The international journal of science / 2 July 2020

nature

OPTICAL PROFUSION

Branched flow of light observed for the first time

Police brutality
How data can help shape policies to combat racial bias

Sustainability
Europe experiences marked rise in forest harvesting

Heat balance
Excitatory neurons that help control body temperature in mice

Vol. 583, No. 7814
nature.com

The international journal of science / 2 July 2020

nature

Is a single target the best way to cut biodiversity loss?

A proposal to limit species extinctions around the world to 'well below' 20 per year needs to be thoroughly assessed.

NICHOLE SOBECKI/*THE WASHINGTON POST*/GETTY

Northern white rhinos have been driven to the brink of extinction.

Next year, all eyes will be on Kunming, China, as talks resume on a new set of global goals to protect biodiversity. These are much needed, because most of the existing 20 targets, which were set in 2010 in Aichi, Japan, have failed to make an impact on the rate of biodiversity loss.

Last month, a team of researchers proposed creating one headline number, suggesting that countries should aim to keep extinctions to "well below" 20 known species every year worldwide (M. D. A. Rounsevell *et al. Science* **368**, 1193–1195; 2020). This would be the biodiversity equivalent of the 2 °C climate target: a simple, measurable goal that can be understood by the public and politicians alike.

The proposal, by Mark Rounsevell at the Karlsruhe Institute of Technology in Germany and his colleagues, is intended to break nearly two decades of failure in global biodiversity policy and target setting — the 2010 Aichi targets replaced a previous unsuccessful target to slow the rate of biodiversity loss that countries set themselves in 2002. And the idea is gaining traction.

In an interview with *Nature*, Elizabeth Maruma Mrema, the new head of the United Nations Convention on Biodiversity, acknowledged that it would be difficult to set a single target because biodiversity is multifaceted. But, if the community succeeds in making it work, she adds: "that will be the best result possible because then it becomes a song everyone will sing, and that everybody can align with to deliver that one key message."

A target for limiting extinctions is not a new idea, and deserves serious consideration. Its feasibility and consequences should be rigorously assessed by the convention's own scientific advisory body, and by the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES), in the same way that climate metrics are assessed by the UN's climate-science advisers, including the Intergovernmental Panel on Climate Change (IPCC).

There are many questions for researchers working in biodiversity to explore. For example, how does a target of 20 extinctions per year — across all plants, animals and fungi — fit with IPBES's own assessment of biodiversity, which says that some one million species are at risk of extinction? Twenty extinctions per year — out of almost two million known species — is ten times higher than the background extinction rate of two per year that existed before humans made a notable contribution to extinctions. But it is considerably lower than today's estimates of species extinctions, which are in excess of 1,000 times the background rate.

> **"The creation of one extinction target should not detract from the need to create nationally relevant targets and policies."**

Other questions include how to choose which species to conserve, and who should make such choices. Would a single number give equal weight to all threatened species, or should those species that are more important to livelihoods and to ecosystem function be given priority for protection? As the authors point out, it is possible for biodiversity loss to result in large and damaging changes to life on Earth without any species going extinct. And at what point would an extinction be declared, given that there is often a time lag between a species going extinct and its being recorded as extinct in the Red List maintained by the International Union for Conservation of Nature?

Given that IPBES's lower estimate for as-yet unidentified plant and animal species is 8.1 million, what are the implications for species that have not yet been described? If policymakers focus resources on conserving known species, what risks might there be to species in parts of the world — such as the marine environment — where knowledge of biodiversity is weak, and which face continued unsustainable development?

And what would the implications of a single target be for the convention's other objectives? Conserving species is one of three aims, alongside ensuring that biodiversity is used sustainably and ensuring that benefits (such as commercial products) are shared fairly, so that no one — for example, Indigenous communities — is left out.

Biodiversity is essential to economic prosperity, food and human health, and the researchers are keen to stress that the creation of one extinction target should not detract from the need for governments to create nationally relevant targets and policies. They also advocate the provision of funding to help countries that are financially poor but biodiversity-rich to meet their goals.

Certainly, a single target, such as that for climate change, would be simpler to communicate than the Aichi targets. And the authors are right to acknowledge

nature

that, ultimately, biodiversity loss continues because public-policy decisions — for example, decisions that lead to industrial economic growth — have not accounted for the costs of replacing the services that species and ecosystems provide to humans.

But they will also know that, although the target to keep global temperatures to within 2 °C of pre-industrial levels was agreed by members of the UN climate convention, that number was subjected to a thorough process of research evaluation by a wide group of researchers in the IPCC before it was adopted.

Any proposal to consider a single numerical target for biodiversity needs to be similarly assessed. IPBES — working with the UN biodiversity convention's own scientific advisers — should be called on to advise. For this to happen, a small group of governments need to make a formal request for scientific advice to the UN convention, and they should do so without delay.

How Europe can fix its forests data gap

The European Union must improve how it collects forest data, which are essential to its ambitions in biodiversity and climate change.

A study published this week reveals how European countries' need for wood biomass is contributing to an increase in forest harvesting (G. Ceccherini *et al. Nature* **583**, 72–77; 2020). The finding comes from a team of researchers at the European Commission's Joint Research Centre in Ispra, Italy, whose conclusions are based on satellite data.

Between the period of 2011–15 and that of 2016–18, 'harvested' forest area — defined as the part of a forest where trees are cut down and others planted in their place — increased by nearly 50%, from 0.76 million hectares to 1.13 million hectares. Of the 26 member states assessed, just 2 — Finland and Sweden — accounted for half of the increase.

This is an important finding. It has implications for biodiversity and climate-change policies, and for the part forests play in nations' efforts to reach net-zero emissions. Forests account for about 38% of the European Union's total land surface, and offset about 10% of its total greenhouse-gas emissions by acting as carbon sinks.

The surge in harvesting might reduce forests' ability to absorb carbon from the atmosphere, the authors say. One reason for this is that large amounts of carbon are released quickly as older trees are felled — but it takes much longer for the same amount of atmospheric carbon to be absorbed by the smaller, younger trees planted in their place.

Paradoxically, the increase in harvested forest area has been driven, in part, by demand for greener fuels, some of which are produced from wood biomass. That includes bioenergy, which comprises about 60% of the EU's renewable energy. This increase in biomass products can, in turn, be traced to the EU's bioeconomy strategy, a policy that has promoted the use of forest resources for energy, as raw materials for industries and to create jobs.

> **"Time is running out. Forest exploitation cannot continue at the current rate."**

The bioeconomy strategy has been a success in one respect: total economic output from the EU's forests between 2012 and 2016 rose by 25%, from €43 billion to €54 billion — and the increase doubled to 50% in Poland and Sweden. But economic success has come at an ecological cost.

Many of the continent's leaders are advocates of a set of ideas known as the European Green Deal, which aims to keep economies growing and create jobs by promoting greener development. However, these objectives can end up counteracting each other. For example, in its new biodiversity strategy, published in May, the EU proposes planting 3 billion trees. But it also suggests designating 30% of land (up from 26%), including old-growth forests, as protected by 2030. If forest harvesting continues at the current rate, such an ambition will be difficult to achieve.

The EU also has a target to double its share of low-carbon and renewable energy to 34% from 2015 to 2030. The European Parliament agreed that the burning of wood could count towards this target. But if wood were to supply even 40% of the extra energy, that would mean burning all of Europe's existing harvest, profoundly threatening the world's forests.

The European Commission is designing a new forestry strategy, expected in 2021, that will complement the biodiversity policy. The Joint Research Centre has been asked by the commission to establish a permanent EU observatory on forests. This will draw on the type of satellite data used in the current study to more regularly monitor deforestation, forest degradation and changes to global forest cover — and will make the data accessible to the public. The researchers drew on data from the joint NASA/US Geological Survey Landsat series of Earth-observation satellites and the Global Forest Change data set, and used Google Earth Engine, a facility that enables researchers to use Google's supercomputers to process satellite imagery.

The planned forest observatory is a crucial development, and one for which the commission deserves to be commended. Once its data become available, EU member states need to incorporate them into the official statistics that policymakers use to make decisions — for example, when planning strategies to reach net-zero emissions. Many countries' forest data — including those that are reported to the EU's statistics office, Eurostat — are based on manual forest surveys. Such surveys are important, but in some cases they are carried out only at decadal intervals, partly because they are expensive. A dedicated observatory will provide decision-makers with much more timely data and help them to identify unintended consequences of their policies.

Ultimately, data must drive action. And, as we have often written, time is running out. Forests provide valuable services on which people and the environment depend. Their exploitation cannot continue at the current rate.

nature

that, ultimately, biodiversity loss continues because public-policy decisions – for example, decisions that lead to industrial economic growth – have not accounted for the costs of replacing the services that species and ecosystems provide to humans.

But they will also know that, although the target to keep global temperatures to within 2 °C of pre-industrial levels was agreed by members of the UN climate convention, that number was subjected to a thorough process of research evaluation by a wide group of researchers in the IPCC before it was adopted.

Any proposal to consider a single numerical target for biodiversity needs to be similarly assessed. IPBES – working with the UN biodiversity convention's own scientific advisers – should be called on to advise. For this to happen, a small group of governments need to make a formal request for scientific advice to the UN convention, and they should do so without delay.

How Europe can fix its forests data gap

The European Union must improve how it collects forest data, which are essential to its ambitions in biodiversity and climate change.

A study published this week reveals how European countries' need for wood biomass is contributing to an increase in forest harvesting (G. Ceccherini *et al. Nature* **583**, 72–77; 2020). The finding comes from a team of researchers at the European Commission's Joint Research Centre in Ispra, Italy, whose conclusions are based on satellite data.

Between the period of 2011–15 and that of 2016–18, 'harvested' forest area – defined as the part of a forest where trees are cut down and others planted in their place – increased by nearly 50%, from 0.76 million hectares to 1.13 million hectares. Of the 26 member states assessed, just 2 – Finland and Sweden – accounted for half of the increase.

This is an important finding. It has implications for biodiversity and climate-change policies, and for the part forests play in nations' efforts to reach net-zero emissions. Forests account for about 38% of the European Union's total land surface, and offset about 10% of its total greenhouse-gas emissions by acting as carbon sinks.

The surge in harvesting might reduce forests' ability to absorb carbon from the atmosphere, the authors say. One reason for this is that large amounts of carbon are released quickly as older trees are felled – but it takes much longer for the same amount of atmospheric carbon to be absorbed by the smaller, younger trees planted in their place.

Paradoxically, the increase in harvested forest area has been driven, in part, by demand for greener fuels, some of which are produced from wood biomass. That includes bioenergy, which comprises about 60% of the EU's renewable energy. This increase in biomass products can, in turn, be traced to the EU's bioeconomy strategy, a policy that has promoted the use of forest resources for energy, as raw materials for industries and to create jobs.

> “Time is running out. Forest exploitation cannot continue at the current rate.”

The bioeconomy strategy has been a success in one respect: total economic output from the EU's forests between 2012 and 2016 rose by 25%, from €43 billion to €54 billion – and the increase doubled to 50% in Poland and Sweden. But economic success has come at an ecological cost.

Many of the continent's leaders are advocates of a set of ideas known as the European Green Deal, which aims to keep economies growing and create jobs by promoting greener development. However, these objectives can end up counteracting each other. For example, in its new biodiversity strategy, published in May, the EU proposes planting 3 billion trees. But it also suggests designating 30% of land (up from 26%), including old-growth forests, as protected by 2030. If forest harvesting continues at the current rate, such an ambition will be difficult to achieve.

The EU also has a target to double its share of low-carbon and renewable energy to 34% from 2015 to 2030. The European Parliament agreed that the burning of wood could count towards this target. But if wood were to supply even 40% of the extra energy, that would mean burning all of Europe's existing harvest, profoundly threatening the world's forests.

The European Commission is designing a new forestry strategy, expected in 2021, that will complement the biodiversity policy. The Joint Research Centre has been asked by the commission to establish a permanent EU observatory on forests. This will draw on the type of satellite data used in the current study to more regularly monitor deforestation, forest degradation and changes to global forest cover – and will make the data accessible to the public. The researchers drew on data from the joint NASA/US Geological Survey Landsat series of Earth-observation satellites and the Global Forest Change data set, and used Google Earth Engine, a facility that enables researchers to use Google's supercomputers to process satellite imagery.

The planned forest observatory is a crucial development, and one for which the commission deserves to be commended. Once its data become available, EU member states need to incorporate them into the official statistics that policymakers use to make decisions – for example, when planning strategies to reach net-zero emissions. Many countries' forest data – including those that are reported to the EU's statistics office, Eurostat – are based on manual forest surveys. Such surveys are important, but in some cases they are carried out only at decadal intervals, partly because they are expensive. A dedicated observatory will provide decision-makers with much more timely data and help them to identify unintended consequences of their policies.

Ultimately, data must drive action. And, as we have often written, time is running out. Forests provide valuable services on which people and the environment depend. Their exploitation cannot continue at the current rate.

World view

By Maharaj K. Pandit

The Himalaya should be a nature reserve

Conservation could be part of the toolkit for diplomacy between China and India.

Thirty-five years ago, at the beginning of my research career, I walked for weeks to study populations of the endangered Himalayan goldthread or *mishmi teeta* (*Coptis teeta*), an endemic plant in Arunachal Pradesh in the Eastern Himalaya that is used as a potent antimalarial drug by local communities. Himalayan species are intriguing. Like isolated islands, mountain peaks reveal how evolution works: by knowing where unique species concentrate, we can learn how speciation occurs.

It would be hard for researchers across Himalaya to do research treks today. Hundreds of thousands of soldiers are now stationed across the high Himalaya. Mid-June saw the worst clash in 45 years, when India–China disengagement talks were followed by a fatal brawl that left at least 20 people dead, several from falls into a river gorge. It is the latest episode in a border conflict between two nuclear powers, and it is happening in a unique, fragile ecosystem.

The Himalaya, which straddles seven nations, already has one of the world's highest rates of deforestation as a result of logging, agricultural expansion, a burgeoning human population, and the building of dams and other infrastructure. It is also thought to be the most rapidly warming mountain range on Earth. Alongside the animal species, Himalayan alpine meadows boast a wealth of herbaceous flowering plants – strange, colourful and delicate – often with medicinal properties. Nowhere else are so many native plant species found at such high elevations.

I have been studying this region for decades, mainly investigating the effects of dams, deforestation, land-use changes, conservation and policy. Roads and buildings to accommodate troops are encroaching on this fragile territory. Pangong Tso Lake, at an altitude of 4,280 metres, saw a military face-off in May. It is only one of many unique Himalayan ecosystems under boots. The lake is a specialized saline water body surrounded by alpine meadows. Militarization, land-use changes, and habitat destruction and fragmentation across the Himalaya are likely to push several species with small populations to extinction. Diplomacy is their only hope.

Here is my idealistic aim for this region. Alongside other multilateral strategies, the mountain range, or at least those areas between 2,600 and 4,600 metres high – whose famous inhabitants include the snow leopard and its prey, the Himalayan blue sheep – should be designated a nature reserve. I propose calling it the Himalaya-one-Nature-one-Reserve, or HONOR. It would ideally encompass much of the Himalayan biodiversity hotspot in the Eastern and Western Himalaya, about 740,000 square kilometres.

> “Some sort of stand-down will happen – and conservation should come into the discussions.”

My dream is not as far-fetched as it sounds. In Antarctica, the Ross Sea Marine Protected Area covers more than 1.5 million square kilometres under a 25-nation agreement. The largest land-based protected area, Northeast Greenland National Park, is 972,000 square kilometres.

I am also inspired by other conservation efforts. The Mekong River Commission includes the governments of Cambodia, Laos, Thailand and Vietnam. A similar Himalayan River Commission, involving all the Himalayan headwater and downstream nations, needs to be explored. In the Himalaya, fledgling transnational conservation efforts and proposals, such as the Kailash Sacred Landscape Conservation and Development Initiative, and the Kangchenjunga Landscape Conservation and Development Initiative, should be ratified and strengthened.

These ideas need to be on the table now, while tensions are so alarmingly high. None of the Himalayan countries wants war, so some sort of stand-down will happen – and conservation should come into the discussions. There is a de-facto code for military engagements at this border to avoid the use of firearms. Surely, not building more infrastructure is as feasible as soldiers not using guns.

The military infrastructure built so far in fragile parts of the Himalaya includes tens of thousands of kilometres of roads. The Chinese-backed US$75-billion China–Pakistan Economic Corridor is a 3,000-kilometre-long route comprising roads, a railway and oil pipelines. India's Border Roads Organisation has been empowered to build 3,400 kilometres of strategic border roads, 61 in total, to cater for far-flung communities, pilgrims and border security.

Transporting fuel to inaccessible terrains to melt bitumen for the road surface is expensive and arduous. Woody plants such as rhododendrons, oaks and conifers, including extremely slow-growing shrubs such as *Juniperus*, are regularly used as fuelwood. A Belt and Road initiative of the Chinese government, which India is not involved in, passes through the most fragile Himalayan landscapes.

As the grasses and herbaceous plants disappear from these alpine valleys, so will a way of life. With no public health-care system, imperilled medicinal herbs are the only source of community medicine, and the only source of cash for highland marginal communities. The semi-domesticated yak in the Himalayan highlands, on which the people depend, cannot graze on the shrubs that are fast invading the meadows under the impact of global warming.

I dream instead of the Himalayan highlands transformed into a peaceful nature reserve, and that the huge public funds squandered on managing conflict are invested instead in infrastructure for health care, education, conservation and welfare. Perhaps this vision will inspire those urgently trying to bring peace to the roof of the world.

Maharaj K. Pandit is director of the Centre for Inter-Disciplinary Studies of Mountain & Hill Environment and at the University of Delhi. He is the author of *Life in the Himalaya: An Ecosystem at Risk*.
e-mail: rajkpandit@gmail.com

World view

By Christopher Monroe

Remote quantum computing is the future

Conditions created by the COVID-19 shutdown are delivering one experiment's 'best data ever'.

The COVID-19 pandemic and shutdown have been disastrous for many people. But one research project in my lab has been humming along, taking the best data my team has ever seen. It is an advanced 'ion trap' quantum computer, which uses laser beams to control an array of floating atoms.

We spent three years setting it up to run remotely and autonomously. Now, we think more labs should run quantum-computing experiments like this, to speed up research.

Quantum computers exploit the weird behaviour of matter at the atomic level. One particle can store many pieces of information, allowing the computers, in effect, to perform many calculations simultaneously. They promise to solve problems that are out of reach of conventional machines, and to speed up modelling of chemical reactions in batteries or drug design, or even simulations of information flow in black holes.

But good quantum hardware is extremely fragile, and the larger the system, the more easily it is perturbed. Some quantum components must be chilled to near absolute zero. Others must be stored in a vacuum more rarefied than that of outer space. It's really hard to prepare and control precise quantum states, let alone keep them stable for hours. Stray currents, changes in temperature and vibrations can easily destabilize the system.

The quantum computer at the University of Maryland, led by myself and physicist Marko Cetina, uses up to 32 identical atoms as the quantum bits, or qubits. Each is levitated by electromagnetic fields and cooled by lasers to sit almost at rest. Typically, such an apparatus has thousands of electronic and optical components, all aligned precisely on a 3-metre wide, 500-kilogram steel table damped against vibrations. It requires an army of people to tweak mirrors and adjust signals, and the components must continually be replaced, tested, calibrated and updated.

But in 2016,we decided to redesign our system to run remotely — not just for convenience, but because that's what our research goals require. We needed to add more qubits without increasing noise and errors, to test complex quantum gate operations, circuits and algorithms.

This required a different approach. For qubits, we use particular states of ytterbium-171 that are so stable that they are widely used for atomic clocks. We miniaturized the most reliable control components, added transducers and feedback circuits, and ran everything from an open-access software platform. We worked closely with many industry partners to make it all work, in a collaboration with engineers Jungsang Kim and Kenneth Brown at Duke University.

> **"It would be great if some quantum components were 'plug-and-play', like the flash memory on a smartphone camera."**

Christopher Monroe is a physicist and engineer at the University of Maryland in College Park and at Duke University in Durham, North Carolina. He is also co-founder and chief scientist at IonQ.
e-mail: monroe@umd.edu

EURIQA (Error-corrected Universal Reconfigurable Ion trap Quantum Archetype) began operating autonomously in April 2019. The whole system now sits in a 1-metre-cubed box. It's rarely opened. One researcher visits the lab for 10–20 minutes once a week to reboot the odd computer that has frozen or power supply that has tripped.

Since my university went into COVID-19 shutdown in March, EURIQA has kept running — all day, every day. And the data have been excellent because the campus has been a ghost town. The lab's temperature hasn't wavered and there's little vibrational noise in the unoccupied building. It's one of very few university quantum experiments making real progress right now.

But there's a bigger picture. This remote mode of operation is exactly what's needed in quantum-computing research. Companies including IBM, Google, Honeywell and a start-up I co-founded, IonQ (whose systems are based on EURIQA), are opening up commercial access to their early quantum-computing devices. By the end of 2020, several types of quantum computer will be available through cloud services hosted by Amazon and Microsoft. But researchers won't have access to the inner workings to advance bespoke designs for particular scientific applications. They won't be able to 'co-design', or fully exploit the interplay between computer fabrication and computer use.

Right now, most quantum-computing research involves the study of qubit properties, quantum gate operations and their control. In some cases, the components are wired together for a specific scientific application. Qubits, quantum logic operations and modes for executing programs are selected and optimized for one purpose. It would be great if instead, some of those components were simple 'plug-and-play' commodities, like the flash memory on a smartphone camera. Then, researchers wouldn't have to build everything from scratch: they could insert a module, tweak a parameter or remotely reprogram a circuit.

Such a system could be built by adapting a particular qubit technology and piling stacks of control hardware and software on top, as we have done with EURIQA. Qubits could be swapped and systems redesigned as technology evolves — just as in conventional computing, the vacuum-tube switches of the 1940s gave way to germanium semiconductors and then silicon wafers in the 1960s.

Large quantum-computing initiatives in the United States, Europe, China, Canada, Australia, Singapore and Russia are investing in qubit research while also giving researchers access to commercial cloud services. But extensive research is needed between these extremes. Industry will ultimately mass-produce quantum computers, but the early 'killer apps' might well come from scientific discovery. Unleashing 'full stack' quantum computers into the research community will hasten that search.

News in brief

SECOND-DEADLIEST EBOLA OUTBREAK EVER ENDS IN DEMOCRATIC REPUBLIC OF THE CONGO

An outbreak of the Ebola virus in the northeastern Democratic Republic of the Congo (DRC) that has been raging since 2018 has officially ended. The World Health Organization (WHO) and the DRC government announced the end on 25 June – 42 days after the last case – but it comes as a fresh Ebola outbreak spreads in the country's northwest.

"We are extremely proud to have emerged victorious over an epidemic that has lasted a long time," said Jean-Jacques Muyembe Tamfum, a co-discoverer of Ebola and director of the National Institute for Biomedical Research in Kinshasa, at a press briefing.

The outbreak was declared in August 2018; the virus infected at least 3,470 people, killing 66% of them. That makes it the world's second-largest outbreak of the haemorrhagic disease, after the 2014–16 West Africa epidemic, which killed more than 11,000 people. Experts also say that the northeastern epidemic – which mainly affected North Kivu and Ituri provinces – was one of the most complex health emergencies the world has ever seen, because it occurred in a region of the DRC plagued by 25 years of war and political instability.

But it was the first Ebola outbreak in which a vaccine for the virus was widely deployed. The vaccine, made by drug company Merck and first tested during the West Africa epidemic, was given to more than 300,000 people who had been in close proximity to people with Ebola, and their contacts. More than 80% of people who were vaccinated didn't end up with the disease, said Muyembe, and those who developed Ebola after vaccination had milder cases. Two antibody-based drugs also showed promise in a clinical trial.

Ebola responders now want to replicate these tools and strategies in Équateur, a province on the opposite side of the country, where 18 people have been reported to be infected with Ebola since an outbreak was declared there on 1 June.

MANY PEOPLE WITH CORONAVIRUS DON'T GET A COUGH OR FEVER

A survey of thousands of people in Italy suggests that a striking share of those infected with the new coronavirus never show classic symptoms of COVID-19. In the study, less than one-third of people infected with SARS-CoV-2 fell ill with respiratory symptoms or fever.

More than 16,000 people have died of COVID-19 in Lombardy, the epicentre of Italy's coronavirus outbreak. Piero Poletti at the Bruno Kessler Foundation in Trento, Italy, Marcello Tirani at the Health Protection Agency of Pavia in Italy and their colleagues studied people in Lombardy who had had close contact with an infected person.

Roughly half of these 5,484 contacts became infected themselves (P. Poletti *et al.* Preprint at https://arxiv.org/abs/2006.08471; 2020). Of those, 31% developed respiratory symptoms – such as a cough – or a fever. Only 26% of those under the age of 60 did so. As a person's age increased, so did their odds of experiencing symptoms and becoming ill enough to require intensive care, or to die.

The findings, which have not yet been peer reviewed, could inform hospitals' outbreak preparations, the authors say.

QUIET STAR IS HOME TO TWO INTRIGUING PLANETS

Astronomers have discovered two planets a little more massive than Earth orbiting a nearby star. Unlike many other stars hosting planetary systems, this one is relatively inactive – so it doesn't emit flares of energy that could hurt the chances of life existing on the planets.

The star, called GJ 887, is just under 3.3 parsecs (10.7 light years) from Earth, in the constellation Piscis Austrinus. It is the brightest red-dwarf star visible from Earth.

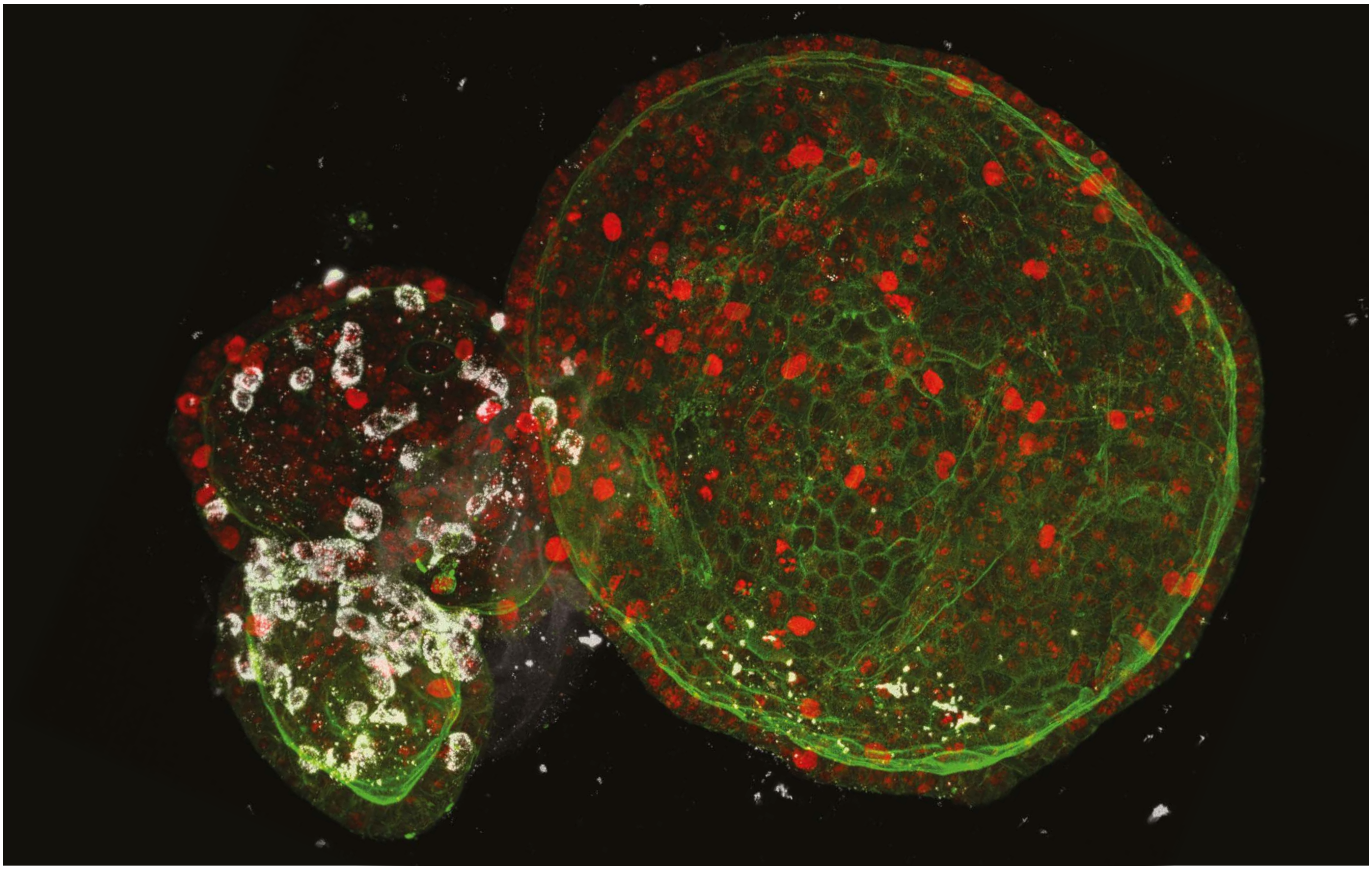
Red dwarfs are smaller and cooler than the Sun, and many have planets orbiting them. But most are very active, with magnetic energy roiling their surface and releasing floods of charged particles into space. Astronomers say the planets in these systems might not be able to support life, because their stars constantly blast them with powerful radiation.

By contrast, planets in the newfound system (artist's impression pictured) could survive relatively unscathed (S. V. Jeffers *et al. Science* **368**, 1477–1481; 2020).

"GJ 887 is exciting because the central star is so quiet," says Sandra Jeffers, an astronomer at Göttingen University in Germany who led the discovery team. "It's the best star in close proximity to the Sun to understand whether its planets have atmospheres and whether they have life."

L TO R: BAZ RATNER/REUTERS; MARK GARLICK

News in focus



JOEP BEUMER/CLEVERS GROUP/HUBRECHT INSTITUTE

Human intestinal organoids infected with SARS-CoV-2 (white).

MINI ORGANS REVEAL HOW THE CORONAVIRUS RAVAGES THE BODY

The virus can damage lung, liver and kidney tissue grown in the lab, which might explain some severe COVID-19 complications in people.

By Smriti Mallapaty

Researchers are growing miniature organs in the laboratory to study how the new coronavirus ravages the body. Studies in these organoids are revealing the virus's versatility at invading organs, from the lungs to the liver, kidneys and gut. Researchers are also testing drugs in these mini tissues to see whether they might be candidates for treating people.

Physicians know from hospitalised patients and autopsies that SARS-CoV-2 can have a devastating effect on organs. But it's unclear whether some of this damage is directly caused by the virus or by complications of the infection. Multiple groups are using organoid studies to show where in the body the virus travels, which cells it infects and what damage it does. "The beauty of organoids is that they resemble the true morphology of tissues," says Thomas Efferth, a cell biologist at Johannes Gutenberg University of Mainz, Germany.

Virologists typically study viruses using cell lines or animal cells cultured in a dish[1]. But these don't model SARS-CoV-2 infection well, say researchers. Organoids better demonstrate what SARS-CoV-2 does to human tissue, says Núria Montserrat, a stem-cell biologist at the Institute for Bioengineering of Catalonia in Barcelona, Spain. They can be grown to include multiple cell types, and they take the shape of the original organ in weeks, she says. They are also less expensive than animal models, and avoid the ethical concerns they pose.

But studies of SARS-CoV-2 in organoids have limitations because they do not reflect the crosstalk between organs that happens in the body. This means that findings will still need to be validated in animal models and clinical studies, says Bart Haagmans, a virologist at Erasmus MC in Rotterdam, the Netherlands.

One of the key insights from organoids is what SARS-CoV-2 does to cells in the respiratory system. Kazuo Takayama, a stem-cell biologist at Kyoto University, Japan, and his colleagues have developed bronchial organoids with four distinct cell types, made from frozen cells from the outer bronchial layer, or epithelium. When they infected the organoids with SARS-CoV-2, they found that the virus mainly targets stem cells that replenish epithelial basal cells, but did not easily enter protective, secretory 'club cells'[2]. The team, which posted its work on bioRxiv, now plans to study whether the virus can spread from basal to other cells.

Respiratory failure

From the upper airways, the virus can enter the lungs and cause respiratory failure, a severe complication of COVID-19. Using mini lungs in a dish, Shuibing Chen, a stem-cell biologist at Weill Cornell Medicine in New York City, has shown that some cells die after being infected, and that the virus induces the production of proteins known as chemokines and cytokines[3], which can trigger a massive immune response. Many people with severe COVID-19 experience this 'cytokine storm', which can be deadly.

But Chen, who also posted her results on bioRxiv, says that why lung cells are dying in patients remains a mystery — whether it's because of damage caused by the virus, self-induced destruction, or through being gobbled up by immune cells. Chen's approach to creating organoids was different from Takayama's: instead of adult cells, she used pluripotent stem cells that can develop into any cell type. Organoids grown in this way can include more cell types, but the final result is less mature and so might not represent adult tissue, says Chen.

From the lungs, SARS-CoV-2 can spread to other organs, but researchers weren't sure how exactly the virus travels until Montserrat and her colleagues published a study in *Cell* in May[4]. In experiments in organoids, also made from pluripotent stem cells, they showed that SARS-CoV-2 can infect the endothelium — the cells lining the blood vessels — which then allows viral particles to leak out into the blood and circulate around the body. Damaged blood vessels in people with COVID-19 also support this hypothesis, says Josef Penninger, a genetic engineer at the University of British Columbia in Vancouver, Canada, and co-lead author of the study[5].

Studies in organoids suggest that once in the blood, the virus can infect several organs including the kidney, say Penninger and Montserrat. Although it infected kidney organoids and some cells died, the researchers are not sure whether this is the direct cause of the kidney dysfunction observed in some people.

Another study in liver organoids found that the virus can infect and kill cholangiocytes — cells that contribute to bile production. Many researchers thought that liver damage seen in COVID-19 was caused by an overactive immune response or drug side effects, says Bing Zhao, a cell biologist at Fudan University in Shanghai, China, who published his results in *Protein & Cell*[6]. His work "suggests that the virus can directly attack the liver tissue, which can cause liver damage", says Zhao.

The virus can also destroy cells that control blood sugar in pancreatic organoids — which adds to mounting evidence that the virus can trigger diabetes in some people (see page 16).

Although such findings are illuminating, using organoids to study the virus–host interaction is in its infancy, says Haagmans, who has studied the virus in gut organoids. "It is too early to say how relevant they are," he says. More complex organoid systems are needed to better understand how the virus interacts with the body's immune system to cause damage, say researchers.

"We are fairly confident now that the virus that causes COVID-19 can infect tissue outside the lung and significantly contribute to disease," says Penninger. But more severe outcomes, such as kidney and heart damage, are probably due to viral infection and an excessive immune response, he says.

Scientists are also studying whether organoids can be used to assess potential COVID-19 therapies, some of which have already been rushed through to clinical trials without extensive testing in cell and animal models. "Due to the time sensitivity, many clinical trials were designed based on previous knowledge of other coronaviruses and launched without careful evaluation in model systems," says Chen. "As a result, many of them have failed."

Chen screened some 1,200 drugs approved by the US Food and Drug Administration for other illnesses, and found that the cancer medication imatinib suppressed SARS-CoV-2 in lung organoids[3]. Several human clinical trials of the drug in treating COVID-19 are under way.

Other groups are also testing existing drugs in organoids, with some success against coronavirus[2,7]. "We will only know at the end of this process what the predictive value of these systems is for testing drug efficacy," says Haagmans. "This is a long-term process."

1. Takayama, K. *Trends Pharmacol. Sci.* https://doi.org/10.1016/j.tips.2020.05.005 (2020).
2. Suzuki, T. *et al.* Preprint at bioRxiv https://doi.org/10.1101/2020.05.25.115600 (2020).
3. Han, Y. *et al.* Preprint at bioRxiv https://doi.org/10.1101/2020.05.05.079095 (2020).
4. Monteil, V. *et al. Cell* **181**, 905–913 (2020).
5. Varga, Z. *et al. Lancet* **395**, 1417–1418 (2020).
6. Zhao, B. *et al. Protein Cell* https://doi.org/10.1007/s13238-020-00718-6 (2020).
7. Krüger, J. *et al.* Preprint at bioRxiv https://doi.org/10.1101/2020.06.10.144816 (2020).

EVIDENCE SUGGESTS THE CORONAVIRUS MIGHT TRIGGER DIABETES

Mounting clues from tissue studies and individuals show the virus can damage insulin-producing cells.

By Smriti Mallapaty

In mid-April, Finn Gnadt, an 18-year-old student from Kiel, Germany, learnt that he had been infected with the SARS-CoV-2 coronavirus despite feeling well. Gnadt's parents had fallen ill after a river cruise in Austria, so his family was tested for virus antibodies, which are produced in response to infection.

Gnadt thought he had endured the infection unscathed, but days later, he started to feel worn out and exceedingly thirsty. In early May, he was diagnosed with type 1 diabetes, and his physician, Tim Hollstein at the University Hospital Schleswig-Holstein in Kiel, suggested that the sudden onset might be linked to the viral infection.

In most people with type 1 diabetes, the body's immune cells start destroying β-cells — which are responsible for producing the hormone insulin — in the pancreas, often suddenly. Hollstein suspected that the virus had destroyed Gnadt's β-cells, because his blood didn't contain the types of immune cell that typically cause the damage.

Diabetes is already known to be a key risk factor for developing severe COVID-19 (ref. 1) and people with the condition are more likely to die from the infection[2]. "Diabetes is dynamite if you get COVID-19," says Paul Zimmet, who studies the metabolic disease at Monash University in Melbourne, Australia.

Now Zimmet is among a growing number of researchers who think that diabetes doesn't just make people more vulnerable to the coronavirus, but that the virus might also trigger diabetes in some[3]. "Diabetes itself is

BERNARD CHANTAL/ALAMY

People with type 1 diabetes, a known COVID-19 risk factor, can't produce the hormone insulin.

a pandemic just like the COVID-19 pandemic. The two pandemics could be clashing," he says.

Their hunch is based on a handful of people such as Gnadt, who have spontaneously developed diabetes[4] after being infected with SARS-CoV-2, and on evidence from dozens more people with COVID-19 who have arrived in hospital with extremely high levels of blood sugar and ketones[5], which are produced from fatty deposits in the liver. When the body doesn't make enough insulin to break down sugar, it uses ketones as an alternative source of fuel.

Researchers cite other evidence, too. Various viruses, including the one that causes severe acute respiratory syndrome (SARS), have been linked with autoimmune conditions such as type 1 diabetes[6]. And many organs involved in controlling blood sugar are rich in a protein called ACE2, which SARS-CoV-2 uses to infect cells[7].

The latest clue comes from an experimental study in miniature lab-grown pancreases. Published last month[8], the work suggests that the virus might trigger diabetes by damaging the cells that control blood sugar.

But other researchers are cautious about such suggestions. "We need to keep an eye on diabetes rates in those with prior COVID-19, and determine if rates go up over and above expected levels," says Naveed Sattar, a metabolic-disease researcher at the University of Glasgow, UK.

To establish a link, researchers need more robust evidence, says Abd Tahrani, a clinician–scientist at the University of Birmingham, UK.

One initiative is now under way. Earlier this month, an international group of scientists, including Zimmet, established a global database[3] to collect information from people with COVID-19 and high blood-sugar levels who do not have a history of diabetes or problems controlling their blood sugar.

Cases are beginning to trickle in, says Stefan Bornstein, a physician at the Technical University of Dresden, Germany, who also helped to establish the registry. The researchers hope to use the cases to understand whether SARS-CoV-2 can induce type 1 diabetes or a new form of the disease. And they want to investigate whether the sudden-onset diabetes becomes permanent in people who've had COVID-19. They also want to know whether the virus can tip people who were already on their way to developing type 2 diabetes into a diabetic state.

The organoid study shows how SARS-CoV-2 could be damaging the pancreas[8]. Shuibing Chen, a stem-cell biologist at Weill Cornell Medicine in New York City, and her colleagues showed that the virus can infect the organoid's α- and β-cells, some of which then die. Whereas β-cells produce insulin to decrease blood-sugar levels, α-cells produce the hormone glucagon, which increases blood sugar. The virus can also induce the production of proteins known as chemokines and cytokines, which can trigger an immune response that might also kill the cells, according to the study, which was published in *Cell Stem Cell* on 19 June.

Chen says the experiments suggest that the virus can disrupt the function of key cells involved in diabetes — by directly killing them or by triggering an immune response that attacks them.

The virus also attacked pancreatic organoids that had been transplanted into mice, and cells in liver organoids. The liver is important for storing and releasing sugar into the blood stream when it senses insulin.

The organoid study adds strength to the argument that SARS-CoV-2 might cause or worsen diabetes, but the paper itself is not enough to prove the link, says Tahrani.

There could be more going on than some scientists suggest, says Shane Grey, an immunologist at the Garvan Institute of Medical Research in Sydney, Australia. The virus could trigger an extreme inflammatory state, which would impair the ability of the pancreas to sense glucose and release insulin, and dampen the ability of the liver and muscles to detect the hormone, he says. This could trigger diabetes.

Only long-term studies will reveal what's really going on, says Sattar.

1. Docherty, A. B. *et al*. Preprint at medRxiv https://doi.org/10.1101/2020.04.23.20076042 (2020).
2. Zhu, L. *et al. Cell Metab.* **31**, 1068–1077 (2020).
3. Rubino, F. *et al. N. Engl. J. Med.* https://doi.org/10.1056/NEJMc2018688 (2020).
4. Chee, Y. J., Ng, S. J. H. & Yeoha, E. *Diabetes Res. Clin. Pract.* **164**, 108166 (2020).
5. Li, J. *et al. Diabetes Obes. Metab.* https://doi.org/10.1111/dom.14057 (2020).
6. Yang, J.-K., Lin, S.-S., Ji, X.-J. & Guo, L.-M. *Acta Diabetol.* **47**, 193–199 (2010).
7. Hamming, I. *et al. J. Pathol.* **203**, 631–637 (2004).
8. Yang, L. *et al. Cell Stem Cell* https://doi.org/10.1016/j.stem.2020.06.015 (2020).

CRISPR EDITING WREAKS CHROMOSOMAL MAYHEM IN HUMAN EMBRYOS

Studies showing large DNA deletions and reshuffling heighten concerns about heritable genome editing.

By Heidi Ledford

A suite of experiments that use the gene-editing tool CRISPR–Cas9 to modify human embryos have revealed that the process can make large, unwanted changes to the genome at or near the target site.

The studies were published last month on the preprint server bioRxiv, and have not yet been peer-reviewed[1–3]. But taken together, they give scientists a good look at what some say is an underappreciated risk of CRISPR–Cas9 editing. Previous experiments have revealed that the tool can make 'off target' gene mutations far from the target site, but the nearby changes identified in the latest studies can be missed by standard assessment methods.

"The on-target effects are more important and would be much more difficult to eliminate," says Gaétan Burgio, a geneticist at the Australian National University in Canberra.

These safety concerns are likely to inform the ongoing debate over whether scientists should edit human embryos to prevent genetic diseases — a process that is controversial because it makes a permanent change to the genome that can be passed down for generations. The first laboratory experiments using CRISPR to edit human embryos took place in 2015. But such studies are still rare and are generally strictly regulated. When, in 2018, biophysicist He Jiankui — the only person known to have edited human embryos that were used for reproduction — revealed the birth in China of twin babies with edited genomes, the work was widely condemned as unethical. He has since been given a prison sentence for "illegal medical practice".

"If human embryo editing for reproductive purposes, or germline editing, were space flight, the new data are the equivalent of having the rocket explode at the launch pad before take-off," says Fyodor Urnov, who studies genome editing at the University of California, Berkeley, but was not involved the latest works.

Unwanted effects

The current research underscores how little is known about how human embryos repair DNA cut by the genome-editing tools — a key step in CRISPR–Cas9 editing — says reproductive biologist Mary Herbert at Newcastle University, UK. "We need a basic road map of what's going on in there before we start hitting it with DNA-cutting enzymes," she says.

The first preprint was posted online on 5 June by developmental biologist Kathy Niakan at the Francis Crick Institute in London and her colleagues. In that study[1], the researchers used CRISPR–Cas9 to create mutations in the *POU5F1* gene, which is important for embryonic development. Of 18 genome-edited embryos, about 22% contained unwanted changes affecting large swathes of the DNA surrounding *POU5F1*. These included DNA rearrangements and large deletions of several thousand DNA bases — much greater changes than are typically intended.

Another group, led by stem-cell biologist Dieter Egli at Columbia University in New York City, studied[2] embryos created with sperm carrying a blindness-causing mutation in a gene called *EYS*. The team used CRISPR–Cas9 to break the DNA in the *EYS* gene, and found that about half of the embryos lost large segments of the chromosome on which *EYS* is situated — and sometimes all of it.

> "This is something that all of us in the scientific community will take more seriously."

And a third group, led by reproductive biologist Shoukhrat Mitalipov at Oregon Health & Science University in Portland, studied embryos made using sperm with a mutation that causes a heart condition[3]. This team also found signs that editing affected large regions of the chromosome containing the mutated gene.

In all the studies, researchers used the embryos for scientific purposes only, and not to generate pregnancies. The lead authors of the three preprints declined to discuss the details of their work with *Nature*'s news team until the articles are published in peer-reviewed journals.

Editing human embryos is controversial because it makes heritable changes to the genome.

PASCAL GOETGHELUCK/SPL

The changes are the result of DNA-repair processes harnessed by genome-editing tools. CRISPR–Cas9 uses a strand of RNA to direct the Cas9 enzyme to a site in the genome with a similar sequence. The enzyme then cuts both strands of DNA at that site, and the cell's repair systems heal the gap.

The edits occur during that repair process: most often, the cell seals up the cut using an error-prone mechanism that can insert or delete a small number of DNA letters. If researchers provide a DNA template, the cell might use that sequence to mend the cut, resulting in a true rewrite. But broken DNA can also cause shuffling or loss of a large region of the chromosome.

Previous work using CRISPR in mouse embryos and other kinds of human cell has demonstrated that editing genes can cause large, unwanted effects[4,5]. But it was important to demonstrate the work in human embryos, says Urnov, because various cell types might respond to genome editing differently.

Such rearrangements could easily be missed: many experiments look for other unwanted edits, such as single DNA-letter changes or insertions or deletions of only a few letters. But the latest studies looked specifically for large changes near the target. "This is something that all of us in the scientific community will, starting immediately, take more seriously than we already have," says Urnov. "This is not a one-time fluke."

Genetic changes

The three studies offered different explanations for how the DNA changes arose. Egli and Niakan's teams attributed the bulk of the changes observed in their embryos to large deletions and rearrangements. Mitalipov's group instead said that up to 40% of the changes it found were caused by a phenomenon called gene conversion, in which DNA-repair processes copy a sequence from one chromosome in a pair to heal the other.

Mitalipov and his colleagues reported[6] similar findings in 2017, but some researchers were sceptical that frequent gene conversions could occur in embryos. Egli and his colleagues tested for gene conversions in their latest work and didn't find them, and Burgio points out that the assays used in Mitalipov's study are similar to those the team used in 2017. One possibility is that DNA breaks heal differently at various positions along the chromosome, says Jin-Soo Kim, a geneticist at the Institute for Basic Science in Seoul and a co-author of the Mitalipov preprint.

1. Alanis-Lobato, G. *et al.* Preprint at bioRxiv https://doi.org/10.1101/2020.06.05.135913 (2020).
2. Zuccaro, M. V. *et al.* Preprint at bioRxiv https://doi.org/10.1101/2020.06.17.149237 (2020).
3. Liang, D. *et al.* Preprint at bioRxiv https://doi.org/10.1101/2020.06.19.162214 (2020).
4. Adikusuma, F. *et al. Nature* **560**, E8–E9 (2018).
5. Kosicki, M., Tomberg, K. & Bradley, A. *Nature Biotechnol.* **36**, 765–771 (2018).
6. Ma, H. *et al. Nature* **548**, 413–419 (2017).


SAUL LOEB/AFP VIA GETTY

US President Donald Trump has issued new immigration restrictions.

TRUMP TO SUSPEND NEW VISAS FOR FOREIGN SCHOLARS

Latest action sows anxiety and confusion across the scientific workforce.

By Nidhi Subbaraman and Alexandra Witze

With a proclamation issued on 22 June, US President Donald Trump extended and expanded immigration restrictions to limit the entry of foreign workers to the United States. The move set off ripples of alarm among scientists and drew fire from experts concerned about the future of US science.

According to the order, the United States will stop issuing certain categories of foreign-worker visa — notably, the H-1B visa given to foreign faculty members hired at universities and employees hired by tech firms — until the end of the year. The Trump administration characterized the decision as a plan to stave off the economic impact of the coronavirus pandemic, and to prioritize jobs for US citizens.

The freeze, which went into effect on 24 June, will not apply to people who are currently in the United States, or those with other valid documents for entering the country. It provides exemptions for some foreign workers — academics on J-1 visas, often postdoctoral researchers, should be clear, according to a senior administration official. Officers issuing visas at US consulates abroad will evaluate petitions for other exemptions, including requests from researchers or doctors engaged in COVID-19 work.

Experts slammed the move, and argued that foreign talent is necessary to keep the US scientific enterprise competitive.

“This is a huge deal,” says Julia Phillips, a member of the US National Science Foundation’s governing board and former chief technology officer at Sandia National Laboratories in Albuquerque, New Mexico. Last year, the United States issued more than 188,000 H-1B visas across all sectors, according to the Department of State. A January report from the National Science Foundation said that 30% of people in science and engineering jobs in the United States were born outside the country.

Innovation under threat

“We find it extremely concerning, particularly as medical residents are brought in on H-1B visas, and faculty who are necessary to educate the US workforce,” says Lizbet Boroughs, associate vice-president for federal relations at the Association of American Universities in Washington DC, whose members include leading US research institutions.

“The bottom line is that suspending processing for H-1B visas is going to have an impact on American research and American innovation and America’s ability to train and teach its scientific-workforce pipeline,” she says.

For students, postdoctoral fellows and faculty members from overseas, the move adds fresh uncertainty and anxiety to an already tumultuous 2020. In April, citing economic damage from the pandemic, the administration paused the issuing of permanent-residency permits, or green cards, to people outside the United States, although it exempted medical workers. The new order extends the suspension to the end of the year, and adds new categories of visa that will be restricted.

“A lot of people are trying to figure out what this means, how they are going to be personally affected,” says Mehmet Doğan, a Turkish physicist at the University of California, Berkeley, who is part of an immigration working group at the University of California Union of Postdocs and Academic Researchers. He is awaiting an H-1B visa, but with the new rules, the path ahead is unclear.

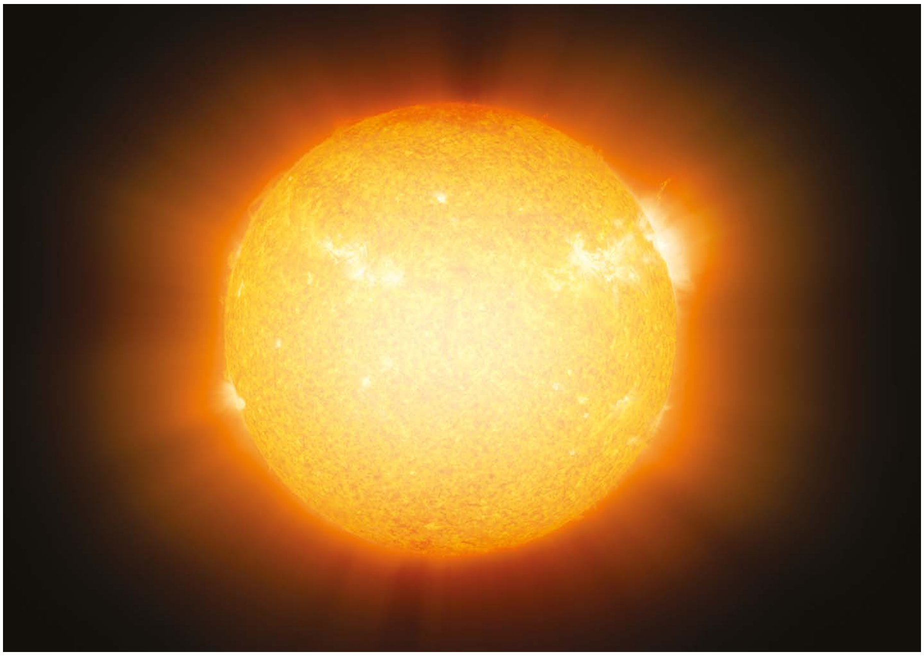
“It is really sad that when this country has so many of the greatest research institutions in the world, greatest universities in the world, that when something like a pandemic happens, one of the first things the government does is to blame international researchers for unemployment,” Doğan says. “That’s crazy, but it’s also very sad.”

‘Limbo is a good term’

Lewis Bartlett, an infectious-disease ecologist at the University of Georgia in Athens, is among those trying to sort out his future. A UK citizen, he applied earlier this year for an H-1B visa to continue his work on the ecology and evolution of infectious diseases in agriculture, particularly to support US beekeeping. He is hoping to have his application approved before his current immigration approval expires. But the executive order has thrown the whole process — already delayed by the pandemic — into question. “There is a lot of uncertainty,” he says. “Limbo is a good term.”

The string of changes to immigration regulations is taking a toll on the students and postdocs who work in the laboratory of pancreatic-cancer researcher Anirban Maitra at the University of Texas MD Anderson Cancer Center in Houston. Most of the 14 people in the group come from outside the United States. “Every day there’s a new rule,” says Maitra. “It’s just continuous stress.”

How the rules will be applied once consulates open after pandemic-associated closures and begin processing visa applications remains to be seen. In the meantime, the new measures send a clear message, says Phillips. “You may be the most brilliant student anywhere. If you were not born in the US, there are absolutely no guarantees whether you will have any option to remain.”



Neutrinos are released during nuclear-fusion reactions in the Sun's centre.

DETLEV VAN RAVENSWAAY/SPL

NEUTRINOS REVEAL FINAL SECRET OF SUN'S NUCLEAR FUSION

Detection of particles produced in the core supports long-held theory about how our star is powered.

By Davide Castelvecchi

By catching neutrinos emanating from the Sun's core, physicists have filled in the last missing detail of how nuclear fusion powers the star.

The detection confirms decades-old theoretical predictions that some of the Sun's energy is made by a chain of reactions involving carbon and nitrogen nuclei. This process fuses four protons to form a helium nucleus, which releases two neutrinos – the lightest known elementary particles of matter – as well as other subatomic particles and copious amounts of energy. This carbon–nitrogen (CN) reaction is not the Sun's only fusion pathway: it produces less than 1% of the Sun's energy. But it is thought to be the dominant energy source in larger stars. The results mark the first direct detection of neutrinos from this process.

"It's intellectually beautiful to actually confirm one of the fundamental predictions of stellar structure theory," says Marc Pinsonneault, an astrophysicist at Ohio State University in Columbus.

The findings, which have not yet been peer reviewed, were reported on 23 June by the Borexino underground experiment in central Italy, at the virtual Neutrino 2020 conference.

The facility was the first to directly detect neutrinos from three distinct steps of a separate reaction, called the proton–proton chain, which accounts for most of the Sun's fusion[1–3]. "With this outcome, Borexino has completely unravelled the two processes powering the Sun," said Borexino co-spokesperson Gioacchino Ranucci, a physicist at the University of Milan, Italy, who presented the results.

> **"It's intellectually beautiful to actually confirm one of the fundamental predictions of stellar structure theory."**

The findings are a final milestone for Borexino, which might now shut down within a year. "We ended with a bang," says the experiment's other co-spokesperson, Marco Pallavicini, a physicist at the University of Genoa, Italy.

Balloon detector

The Borexino solar-neutrino experiment occupies a hall under more than one kilometre of rock in the Gran Sasso National Laboratories near L'Aquila, Italy, where it has been in operation since 2007. The detector is a giant nylon balloon filled with 278 tonnes of liquid hydrocarbons and immersed in water. Almost all neutrinos from the Sun zip through Earth – and Borexino – in a straight line, but a tiny number bounce off electrons in the hydrocarbons, producing flashes of light that are picked up by photon sensors in the water tank.

Because the CN reaction chain is responsible for only a small fraction of solar fusion, neutrinos from it are relatively rare. Moreover, the CN neutrinos are easy to confuse with those produced by the radioactive decay of bismuth-210, an isotope that leaks from the balloon's nylon into the hydrocarbon mixture.

Although the contamination is extremely low – at most, a few dozen bismuth nuclei decay per day inside Borexino – separating the solar signal from bismuth noise required a painstaking effort that began in 2014. The bismuth-210 couldn't be prevented from leaking out of the balloon, so the goal was to slow the rate at which the element seeped into the middle of the fluid, while ignoring signals from the outer edge. To do this, the team had to control any temperature imbalances across the tank, which would produce convection and mix its contents faster. "The liquid must be extraordinarily still, moving at most at a few tenths of centimetres per month," Pallavicini says.

To keep the hydrocarbons at a constant, uniform temperature, the researchers wrapped the entire tank in an insulating blanket and installed heat exchangers to automatically balance the temperature throughout. Then, they waited. It was only in 2019 that the bismuth noise became quiet enough for the neutrino signal to stand out. By early 2020, the researchers had gathered enough of the particles to definitively claim they had detected neutrinos from the CN nuclear-fusion chain.

"It is the first really direct evidence that hydrogen burning through CN operates in stars," says Aldo Serenelli, an astrophysicist at the Institute of Space Sciences in Barcelona, Spain. "So this is really amazing."

Sun-surface speculation

As well as confirming theoretical predictions about what powers the Sun, the detection of CN neutrinos could shed light on the structure of its core – specifically, the concentrations of elements astrophysicists call metals (anything heavier than hydrogen and helium).

The amounts of neutrinos seen by Borexino seem consistent with the standard models in which the 'metallicity' of the Sun's core is similar to that of its surface. But more up-to-date studies have begun to question that assumption, Serenelli says.

These studies suggest that the metallicity is lower. And because these elements regulate how fast heat diffuses from the Sun's core, it implies that the core is slightly cooler than previously estimated. Neutrino production is extremely sensitive to temperature and, taken together, the various amounts of neutrinos seen by Borexino seem to be consistent with the older metallicity values — not with the new ones, Serenelli says.

As a possible explanation, he and other astrophysicists have suggested that the core has higher metallicity than have the outer layers. Its composition could reveal more about early stages of the Sun's life, before the formation of the planets removed some of the metals that were accreting onto the young star.

1. Bellini, G. *et al. Phys. Rev. Lett.* **107**, 141302 (2011).
2. Bellini, G. *et al. Phys. Rev. Lett.* **108**, 051302 (2012).
3. Bellini, G. *et al. Nature* **512**, 383–386 (2014).

NIH SEXUAL-HARASSMENT RULES ARE STILL TOO WEAK, SAY CRITICS

US biomedical research agency has a new policy, but relies on universities to report bad behaviour.

By Nidhi Subbaraman

The US National Institutes of Health (NIH) last week published new guidelines for tracking sexual-harassment complaints involving scientists funded by the agency. On 24 June, it described the actions it will take when alerted to reports of unsafe behaviour, including restricting scientists from peer-review panels, holding back pending grants and refusing university requests to transfer funding to other institutions in cases where a harasser changes jobs.

Advocates who have campaigned for changes at the US$41-billion biomedical-research agency say the adjustments are necessary, but are still weaker than rules issued by other funding agencies, such as the National Science Foundation (NSF).

Measures introduced on 11 June say that universities must inform the NIH when major changes are made to a grant owing to an investigation about scientists creating an unsafe work environment. "We have specifically defined that as including harassment, bullying, sexual harassment and other inappropriate behaviour," says Carrie Wolinetz, NIH associate director for science policy.

The NIH began collecting information about sexual-harassment investigations at the institutions it funds in 2019. But until the June announcement, disclosures had been voluntary. According to NIH officials, the new measures put harassment on the same level as research misconduct, fraud, issues of foreign influence and violations of peer-review integrity.

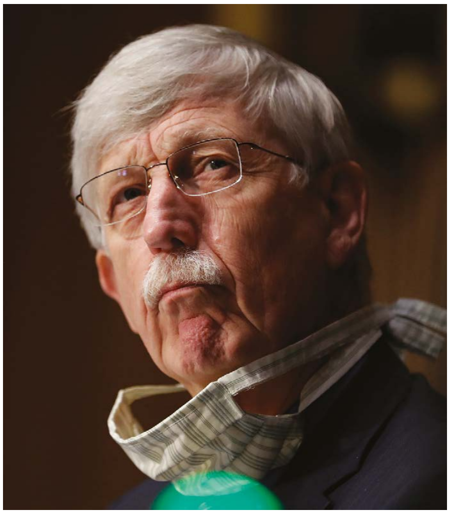
Critics say that the policy still relies too heavily on universities, which might be disinclined to report bad behaviour to the agency that funds them, and that a raft of steps must follow to change the status quo.

It "assumes good faith on the part of the institutions", says BethAnn McLaughlin, a neuroscientist and founder of the non-profit group MeTooSTEM. "What an absurd and insulting notion."

Others are awaiting the agency's next move. "This guidance is a good start, but there is much more that needs to be done," says Angela Rasmussen, a virologist at Columbia University in New York City, who was part of a working group convened by the NIH to examine its policies and suggest ways the agency could improve.

Changes and challenges

Agencies and institutions in the United States have begun making changes after acknowledging the scope and harm of sexual harassment in science. A 2018 report by the US National Academies of Sciences, Engineering, and Medicine in Washington DC found that incidents of harassment are rampant, that such behaviour pushes talented researchers out of science, and that university and federal policies for keeping it in check are lacking.



ANDREW HARNIK/GETTY

NIH director Francis Collins has been criticized for not moving faster to strengthen the agency's policies against harassment.

In a June presentation to a panel of advisers to the NIH director, Wolinetz said that as of 8 June, the NIH had received information about 115 cases of sexual harassment in 2019 and 27 cases in 2020, from 71 institutions. So far this year, it has removed 24 people from peer-review committees. In 2019, it removed 64.

According to the information provided to the NIH, only 14 principal investigators have been removed from grants so far, in part because investigations at their institutions are ongoing. But even in cases in which there have been findings of harassment, some institutions have pushed back against removing the harassers, arguing to keep the funding in place after the offender has been disciplined. "We are starting to see people, upsettingly, try to game the system a little bit," Wolinetz says.

Alysha Dicke, a member of the NIH's working group, is concerned that this pattern will continue if the NIH is not more transparent about affected universities and grants, and about what constitutes reportable behaviour. "I think it's important for NIH to point out how institutions are not responding as intended/desired, as it will likely be even more difficult to change some of the undesirable institutional behaviour if it's never called out," she wrote in an e-mail to *Nature*.

The new guidance won't provide a comprehensive view of harassment at funded institutions. The NIH requests that universities report "concerns" about scientists that have led to changes in grants — including pending investigations. But lawyer Kristina Larsen is sceptical that many institutions will report anything other than the findings of completed investigations — which only rarely occur. Larsen was an administrator at the University of California, San Diego, before she began representing people who filed sexual-harassment complaints. "I don't think it's realistic," she says.

Other funding agencies in the United States have stronger rules. In 2018, the NSF began requiring universities that find that an agency-funded scientist has committed sexual harassment to report this to the NSF within ten business days. NASA adopted similar rules this March. But the NIH rules require reporting only when the status of a grant changes. Wolinetz says that's because the NIH does not have the authority to ask institutions to report investigations or their results outside the grant-update cycle.

"The NSF has direct oversight of civil-rights violations at NSF-funded organizations, and NIH does not," Wolinetz says. "It does present some legal limitations in what we're able to do."

ELIJAH NOUVELAGE/GETTY

Body-worn cameras can increase the accountability of the police, but studies on their use have produced mixed results.

BRUTALITY AND RACIAL BIAS: WHAT THE DATA SAY

Some interventions could help to reduce racism and rein in the use of unnecessary force in police work, but the evidence base is still evolving. **By Lynne Peeples**

For 8 minutes and 46 seconds, Derek Chauvin pressed his knee into the neck of George Floyd, an unarmed Black man. This deadly use of force by the now-former Minneapolis police officer has reinvigorated a very public debate about police brutality and racism.

As protests have spread around the globe, the pressure is on police departments and politicians, particularly in the United States, to do something – from reforming law-enforcement tactics to defunding or even abolishing police departments.

And although researchers are encouraged by the momentum for change, some are also concerned that, without ample evidence to support new policies, leaders might miss the mark. Many have been arguing for years about the need for better data on the use of force, and for rigorous studies that test interventions such as training on how to de-escalate interactions or mandating the use of body-worn cameras. Those data and studies have begun to materialize, spurred by protests in 2014 after the deadly shooting of Michael Brown in Ferguson, Missouri, and the death by chokehold of Eric Garner in New York City.

From these growing data sets come some disturbing findings. About 1,000 civilians are killed each year by law-enforcement officers in the United States. By one estimate, Black men are 2.5 times more likely than white men to be killed by police during their lifetime[1]. And in another study, Black people who were fatally shot by police seemed to be twice as likely as white people to be unarmed[2].

"We have enough evidence that tells us that action needs to be taken," says Justin Nix, a criminologist at the University of Nebraska Omaha. "One thousand deaths a year does not have to be normal." New evidence continues to support a link between racial bias and the use of force. Data from California show that, in 2018, police stopped and used force against Black people disproportionately (see go.nature.com/2bgfrah). A December 2019 paper reported that bias in police administrative records results in many studies underestimating levels of racial bias in policing, or even masking discrimination entirely[3].

The data are still limited, which makes crafting policy difficult. A national data set established by the FBI in 2019, for example, contains data from only about 40% of US law-enforcement officers. Data submission by officers and agencies is voluntary, which many researchers see as part of the problem.

"Most agencies do not collect that data in a systematic way," says Tracey Meares, founding director of the Justice Collaboratory at Yale Law School in New Haven, Connecticut. “I hope when people think about the science of this that they understand what we know, what we don’t know and why we don’t know it,” she says. “Policing, in large part for historical reasons, has proceeded in kind of a science-free zone.”

Bad apples

Scientists must often work around the limitations in the data. Mark Hoekstra, an economist at Texas A&M University in College Station, has attempted to decipher the role of race in police officers’ use of force, by comparing responses to emergency calls.

Based on information from more than two million 911 calls in two US cities, he concluded that white officers dispatched to Black neighbourhoods fired their guns five times as often as Black officers dispatched for similar calls to the same neighbourhoods[4] (see ‘Answering the call’).

Scientists have tried to identify some predictive factors, such as racial bias, a bad temper, insecure masculinity and other individual characteristics, many of which can be identified through simulations already used in officer training[5]. Nix suggests that such screening could help with vetting officers before they are recruited. But raising the bar for hiring might be impractical, he cautions, because many police departments are already struggling to attract and retain highly qualified candidates.

Similar forecasting models could recognize patterns of bad behaviour among officers. Data from the New York City Police Department suggests that officers who had repeated negative marks in their files were more than three times as likely to fire their gun as were other officers[6].

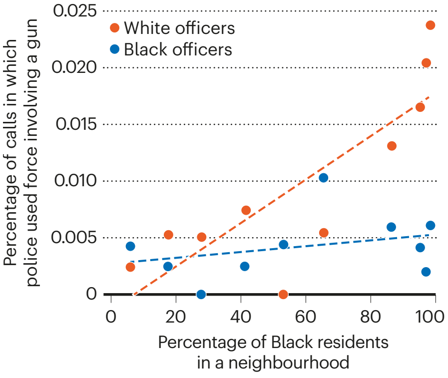
Such wrongdoing might even be contagious. Another study, published in February, looked at complaints filed against police officers in Chicago, Illinois. It found that although only a small percentage of officers shoot at civilians, those who have done so often serve as “brokers” in the social networks within policing[7]. Other officers connected to them were also found to be at greater risk of shooting.

But carrying out disciplinary action, let alone firing a police officer, is notoriously difficult in the United States. Union contracts give officers protections that have been tied to increases in misconduct[8]. In many states, a bill of rights for law-enforcement officers shields personnel from investigations into misconduct. “One thing we need to take a hard look at are those state laws and union contracts that provide either flawed or overly protective procedures that insulate officers from appropriate accountability,” says Seth Stoughton, a former police officer who is a law professor at the University of South Carolina in Columbia.

Lawrence Sherman, director of the Cambridge Centre for Evidence-Based Policing in Cambridge, UK, suggests that states have the constitutional power to license, or revoke, the power of any individual to serve as a police officer. “If a state agency was keeping track of everyone’s disciplinary history, they might have taken Derek Chauvin out of the policing business ten years ago,” says Sherman. Chauvin had received 18 complaints against him even before he put his knee on Floyd’s neck. “We monitor performance of doctors,” Sherman adds. “Why don’t we monitor the performance of police officers?”

ANSWERING THE CALL

Researchers looked at responses to 1.2 million 911 emergency calls in a US city and plotted the use of force involving a gun across neighbourhoods, according to their racial composition. White officers were more likely to use a gun than were Black officers and more likely to do so in predominantly Black neighbourhoods.



SOURCE: REF. 4

Even officers who are fired for misconduct are frequently rehired. The police officer in Cleveland, Ohio, who fatally shot 12-year-old Tamir Rice in 2014 had previously resigned from another police department after it had deemed him unfit to serve. The Cleveland police did not review the officer’s personnel file before hiring him, *The New York Times* reported in 2015. An investigation of public records from Florida showed that about 3% of that state’s police force had previously been fired or had resigned in lieu of being dismissed. The study, published in May, found that these officers tended to move to smaller agencies which served a slightly larger proportion of Black residents, but with no significant difference in crime rates[9]. They also appeared to be more likely to commit misconduct in the future compared to officers who had never been fired.

> **“Policing, in large part for historical reasons, has proceeded in kind of a science-free zone.”**

Federal legislation introduced last month targets barriers to good and fair policing. One bill would effectively end the doctrine of qualified immunity, by which courts have largely prevented officers from being successfully sued for abuse of power or misconduct since the mid-1960s (ref. 10). A similar bill proposes a number of measures intended to increase police accountability, training and data collection, including a national police misconduct registry to keep record of when an officer is fired or quits. Although Democrats in Washington DC broadly support the bills, Republicans unveiled a competing, weaker proposal that does not address the issue of qualified immunity. Robin Engel, director of the Center for Police Research and Policy in Cincinnati, Ohio, suggests that the real capacity for change is at the state and local levels. “There’s a collective citizen call to action now to hold political leaders responsible for ensuring that the police are collecting data, releasing data and operating with best practices,” says Engel.

Evidence-based policing

It remains unclear which law-enforcement practices are actually best, largely because of a lack of data and science. “We’re operating in the dark about what are the most effective strategies, tactics and policies to move forward with,” Engel says. Political leaders and activists pushing for change in the United States have widely endorsed body-worn cameras, de-escalation training, implicit-bias training, early intervention systems, the banning of chokeholds, and civilian oversight since the tragedies of 2014. A survey of 47 of the largest US law-enforcement agencies between 2015 and 2017 found that 39% changed their use-of-force policies in 2015–16 and revised their training to incorporate tactics such as de-escalation. Among the agencies surveyed, officer-involved shootings dropped by 21% during the study period[11].

“But as we have seen in the last several weeks – from Minneapolis and from the police response to the protests – there’s a great deal that still has to change in policing,” says Laurie Robinson, a criminologist at George Mason University in Fairfax, Virginia.

Researchers are advocating collection of better data, such as tracking situations in which force was avoided by de-escalation strategies or, when force was used, recording whether it was at a lower level than it might previously have been.

The Oklahoma City Police Department is among agencies working to fill that void. It now collects details on the applicability of each specific de-escalation tactic and technique any time force is used. “Since the implementation of our de-escalation policy, our use-of-force numbers have decreased,” states Megan Morgan, a police sergeant and spokesperson for the department.

The collection of data might itself hold police officers more accountable. In one study, a requirement that officers file a report when they point their guns at people but do not fire

Protests after the death of George Floyd have renewed pressure to reform US policing.

STEVE SANCHEZ/PACIFIC PRESS VIA ZUMA/EYEVINE

was associated with significantly reduced rates of gun death[12].

The use of body-worn cameras could be among the easiest interventions to enhance accountability. The technology gained traction after a randomized experiment published in 2014 compared shifts in which all officers wore cameras all the time with shifts in which they never did[13]. The likelihood of force being used by officers with cameras was roughly half that of officers without cameras. Furthermore, camera-wearing officers received about one-tenth the number of complaints as did officers without cameras.

Results of more-recent studies have been mixed. When the Las Vegas Metropolitan Police Department in Nevada implemented body cameras, it experienced drops in both the rate of complaints and the use of force[14]. But when the Metropolitan Police Department of the District of Columbia did the same, it found no benefits (see go.nature.com/3heuxac). The differences might have to do with policies that allow officers to choose when to turn on their cameras, as well as a lack of controls for situations in which one officer shows up wearing a camera while another does not, notes Sherman. The latter could dilute true differences in the rates of complaints or uses of force.

"It would be a travesty if we got rid of body cams," says Sherman. "They very often help to clarify what happened."

Evidence suggests that encouraging officers to listen to citizens' views before making decisions and to generally demonstrate an interest in working with members of a community can be another effective intervention. A one-day training programme based on these principles of procedural justice was shown to reduce both citizen complaints and use of force by officers in the Chicago Police Department[15].

"If police are to be of service to communities, they need to build trust with communities that are likely to distrust them," says Thomas O'Brien, a researcher at the Social Action Lab at the University of Illinois in Urbana–Champaign. His work suggests that such trust-building requires the police to both acknowledge its role in creating the distrust, as well as apologize for it[16]. Any half-hearted attempts at reconciliation could backfire, he says. Special training can be difficult, however, particularly in smaller jurisdictions, which have been shown to have a higher rate of police shooting civilians[17] (see 'Small-town problems').

In the wake of Floyd's death, many calls for change have gone beyond police reform to defunding police departments – reducing their public funding and reallocating resources to other programmes – or dismantling them altogether. Some researchers caution against fully abolishing police departments. That could have "disastrous consequences", says Engel. "It's better to work within and demand significant and meaningful change, and then hold them accountable for that change."

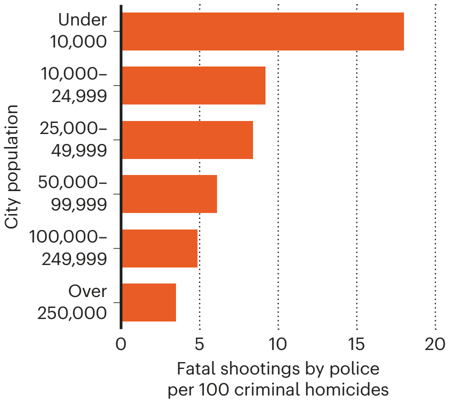
However, Engel does support proposals that would begin "carving off pieces" of law-enforcement agencies' current responsibilities that might fall outside their expertise – or might not require an armed response – such as issues of homelessness, drug abuse and mental illness. In New York City, the police purview goes as far as to include enforcement of street-vendor licences. Across the United States, an arrest is made every 3 seconds; less than 5% of these are for serious violent crimes, according to the Vera Institute of Justice in Brooklyn, New York (see go.nature.com/3fbwmcn).

Curtailing police encounters could also result in fewer crimes. Research published last year found that Black and Latino boys who are stopped more often by police are more likely to commit crimes months later[18].

Stoughton also emphasizes the role of racial bias in society, as evidenced in the months leading up to Floyd's murder – by the fatal shooting of a 25-year-old Black man, Ahmaud Arbery, by two white men while he was jogging in Georgia, and by a white woman's 911 call to falsely report being threatened by a Black birdwatcher in New York City's Central Park. "I have become convinced that we do not have a race problem in policing," says Stoughton. "Rather, we have a race problem in society that is reflected in policing."

SMALL-TOWN PROBLEMS

Large cities account for about 30% of fatal police shootings, but the rate of police shootings per 100 homicides is much higher in smaller communities. Little research has been done to understand this relationship.



SOURCE: REF. 17

Lynne Peeples is a science journalist in Seattle, Washington.

1. Edwards, F., Hedwig, L. & Esposito, M. *Proc. Natl Acad. Sci. USA* **116**, 16793–16798 (2019).
2. Nix, J., Campbell, B. A., Byers, E. H. & Alpert, G. P. *Criminol. Public Policy* **16**, 309–340 (2017).
3. Knox, D., Lowe, W. & Mummolo, J. *Am. Polit. Sci. Rev.* https://doi.org/10.1017/S0003055420000039 (2020).
4. Hoekstra, M. & Sloan, C. W. National Bureau of Economic Research Working Paper 26774 (2020).
5. Goff, P. A. & Rau, H. *Ann. Am. Acad. Polit. Soc. Sci.* **687**, 67–88 (2020).
6. Ridgeway, G. *Stat. Public Policy* **3**, 1–6 (2016).
7. Zhao, L. & Papachristos, A. V. *Ann. Am. Acad. Polit. Soc. Sci.* **687**, 89–112 (2020).
8. Dharmapala, D., McAdams, R. H. & Rappaport, J. University of Chicago Coase-Sandor Institute for Law & Economics Research Paper No. 831 (2018).
9. Grunwald, B. & Rappaport, J. *Yale Law J.* **129**, 6 (2020).
10. Baude, W. 106 California Law Review 45; U of Chicago, Public Law Working Paper No. 610 (2018).
11. Stephens, D. W. *Officer Involved Shootings: Incident Executive Summary* (National Police Foundation, 2019).
12. Jennings, J. T. & Rubado, M. E. *Public Adm. Rev.* **77**, 217–226 (2017).
13. Ariel, B., Farrar, W. A. & Sutherland, A. *J. Quant. Criminol.* **31**, 509–535 (2015).
14. Braga, A., Coldren, J. R., Sousa, W., Rodriguez, D. & Alper, O. *The Benefits of Body-Worn Cameras: New Findings from a Randomized Controlled Trial at the Las Vegas Metropolitan Police Department* (National Criminal Justice Reference Service, 2017).
15. Wood, G., Tyler, T. R., Papachristos, A. V. *Proc. Natl Acad. Sci. USA* **117**, 9815–9821 (2020).
16. O'Brien, T. C., Meares, T. L. & Tyler, T. R. *Ann. Am. Acad. Polit. Soc. Sci.* **687**, 202–215 (2020).
17. Sherman, L. W. *Annu. Rev. Criminol.* **1**, 421–449 (2018).
18. Del Toro, J. *et al. Proc. Natl Acad. Sci. USA* **116**, 8261–8268 (2019).

Books & arts

UPRISING, LLC

Chemist Raychelle Burks shares her experience of gender bias and racial discrimination in the film *Picture a Scientist*.

Three extraordinary women run the gauntlet of science — a documentary

Systemic racism, sexual harassment and institutional bias permeate a film about three female scientists, who have survived and thrived. **Review by Alexandra Witze**

Ask people to picture a scientist, and what do many imagine? A white man in a white lab coat, sadly.

The film *Picture a Scientist* shows why. It chronicles, through the stories of three extraordinary female researchers, the gender and racial biases that drive so many people out of science. All the usual suspects are here: systemic racism, institutional bias, sexual harassment. Together, they tell so many aspiring researchers the lie that they do not belong.

The film-makers interweave interviews with startling statistics. Women receive 50% of the bachelor's degrees in science and technical fields in the United States, yet comprise only 29% of people employed in those fields. The pipeline of people interested in science is full at the start, but it leaks over time because of discrimination and harassment, says Paula Johnson, the president of Wellesley College in Massachusetts.

Implicit bias is pervasive. Men are preferred to women even if they have the same accomplishments. Psychologists have shown this by testing scientists' responses to fictitious CVs that are identical other than coming from 'John' or 'Jennifer', or CVs that include, or scrub, mention of the applicant's status as a member of a minority racial group. Even social scientists who are aware of their own bias do not overcome it,

---

Picture a Scientist
Film by Sharon Shattuck and Ian Cheney/
Uprising Production
Screening online until 26 June at
https://www.pictureascientist.com

Biologist Nancy Hopkins campaigned for equal treatment at work for female scientists.

UPRISING, LLC

as they admit on camera.

The iceberg analogy for sexual harassment is apt. It holds that only a fraction of harassment – obvious things such as sexual assault and sexual coercion – rises into public consciousness and awareness. The rest of the iceberg is buried deep. It includes the more insidious and pernicious attacks, from calling someone horrifying names to sabotaging their lab equipment. "I remember the first time he called me a …" is one of many memorable lines in the film, spoken by a former graduate student of her adviser. And there's a whole other iceberg of covert racial aggression lurking beneath the overt (see, for example, go.nature.com/3hfuco8).

Raychelle Burks has fought harder than most. Burks, an analytical chemist now at the American University in Washington, DC, specializes in developing techniques to detect explosives. We see Burks working in the lab, ebullient in T-shirt and jeans, demonstrating chemistry to students. A Black woman in academia, Burks once got mistaken for a janitor while working at her desk. The higher she rises, the fewer Black scientists there are. Which is why she constantly works in science communication and outreach – many know Burks as Dr Rubidium – so that kids can see a scientist who is a person of colour.

The film-makers follow Burks to a chemistry meeting in Canada, where she talks about diversity to a room of mostly white faces. She tells them that we all code-switch to an extent, changing from our personal to professional personas to interact with other scientists. But no one ever asked, she says, why one version of professionalism – suits, straight hair – is deemed more appropriate than Burks'.

That's as far as *Picture a Scientist* ventures into the intersectional challenges facing many scientists. Its two other protagonists are white women with their own compelling stories.

Biologist Nancy Hopkins was shocked when Francis Crick once put his hands on her breasts as she worked in the laboratory. By the time she became a full professor at the Massachusetts Institute of Technology (MIT) in Cambridge, she knew the problems were both deep-rooted and less obvious. When she couldn't get enough lab space to do her research on zebrafish development, she used a tape measure to prove that male faculty had substantially more space than female faculty. We follow along as Hopkins walks those same hallways today, eyeing the dimensions and tallying up the inequalities.

She recruited colleagues to gather much more data. The culmination was a landmark 1999 study on gender bias in MIT's school of science (see go.nature.com/2ngyiyd), which reverberated across US higher education and forced many administrators to confront entrenched discrimination. Yet Hopkins would rather have spent that time doing science, she relates.

The third story comes from Jane Willenbring, a geoscientist who in 2016 filed a formal complaint accusing her PhD adviser, David Marchant, of routinely abusing her during fieldwork in Antarctica years before. Marchant, who has denied the allegations, was sacked from his post at Boston University in April 2019 after an investigation. *Picture a Scientist* brings Willenbring together with Adam Lewis, who was also a graduate student during that Antarctic field season and witnessed many of the events. Their conversations are a stark reminder of how quickly and how shockingly the filters that should govern work interactions can drop off, especially in remote environments. Lewis tells Willenbring he didn't realize at the time that she had been bothered, because she did not show it. "A ton of feathers is still a ton," she says.

In stark contrast, the film shows us Willenbring, now at the Scripps Institution of Oceanography in San Diego, California, with two of her students working along the coastal cliffs. Slowly, carefully, collaboratively, they drill samples out of the rocks, to extract clues to how California might prove resilient to climate change. It struck me as fitting – given Willenbring's resilience and the strength of the scientists profiled in this film.

Alexandra Witze is a correspondent for *Nature* based in Boulder, Colorado.

Drugs, money and misleading evidence

Take trials out of the hands of drug makers, says a book on corruption in the industry. **By Laura Spinney**

In the race to find treatments and a vaccine for COVID-19, it's more essential than ever that society can trust drug companies seeking regulatory approval. *The Illusion of Evidence-Based Medicine* is the latest in a long line of books that caution us not to hold out much hope.

Child psychiatrist Jon Jureidini and philosopher Leemon McHenry dispute the assumption that all approved drugs and medical devices are safe and effective. They warn that when clinical science is hitched to the pharmaceutical industry's dash for profits, the scientific method is undermined by marketing spin and cherry-picking of data. They propose a solution inspired by philosopher of science Karl Popper: take drug testing out of the hands of manufacturers.

The authors were afraid that academic publishers with ties to the pharmaceutical industry would demand unacceptable changes

ULRICH BAUMGARTEN VIA GETTY

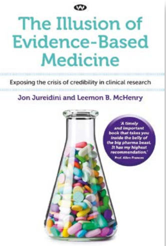
Drug production is a huge industry, with billions of dollars resting on the results of clinical trials.

to their work, so they chose to publish with a small, independent press. To be fair, similar exposés have been produced by mainstream publishers; these include *The Truth About the Drug Companies* (2004) by Marcia Angell, former editor-in-chief of *The New England Journal of Medicine*, and *Bad Pharma* (2012) by the crusading clinical epidemiologist Ben Goldacre.

Little has changed since these works were published, say Jureidini and McHenry. Academics still lend their names to ghost-written papers paid for by drug companies. The companies still pressure journals to publish the papers; on the basis of these, regulators approve drugs. Because the industry controls every aspect of this process – and the all-important data – the pair refer to it as “organized crime”, following Peter Gøtzsche’s 2013 book *Deadly Medicines and Organised Crime*.

Jureidini and McHenry have witnessed these practices at close quarters, and spent more than ten years sifting through documents released by drug companies. In 2007, they were taken on as consultants by a California law firm that has represented plaintiffs in suits against the industry. The duo leave it to readers to decide whether this conflict of interest compromises their position. I am inclined to applaud their determination. “At stake,” they write, “is the integrity of one of the greatest achievements of modern science – evidence-based medicine.”

> **“Distortion of evidence risks further eroding the public’s already fragile trust in academic medicine.”**



The Illusion of Evidence-Based Medicine: Exposing the crisis of credibility in clinical research
Jon Jureidini &
Leemon B. McHenry
Wakefield (2020)

‘Evidence-based medicine’, some might be surprised to learn, was coined as recently as the early 1990s, to highlight the fact that doctors based much of their practice on an unscientific hotchpotch of research, experience, anecdote and custom. It has produced stunning successes, such as treating high blood pressure to reduce the risk of cardiovascular disease, and personalizing the treatment of liver cancer. Yet distortion of evidence threatens those gains, these authors warn, and risks further eroding the public’s already fragile trust in academic medicine, manifesting, for example, in the rising distrust of vaccines.

They discuss two trials for psychiatric drugs: GlaxoSmithKline’s Study 329, testing paroxetine; and Forest Laboratories’ Study CIT-MD-18, testing citalopram. Both aimed to gain US Food and Drug Administration (FDA) approval for the use of antidepressants in children and adolescents. Initial publications concluded that both drugs were safe and effective in that group. Paroxetine was not approved for this use; escitalopram, a variant of citalopram, was.

Analysing the clinical report for Study 329, Jureidini and others found in 2015 that paroxetine was not effective in adolescents with major depression, as the original 2001 publication had claimed. They also found it increased the risk of harms such as suicidal ideation (J. Le Noury *et al. Br. Med. J.* **351**, h4320; 2015). A year later, Jureidini and McHenry deconstructed Study CIT-MD-18 (J. N. Jureidini *et al. Int. J. Risk Safety Med.* **28**, 33–43; 2016). They revealed that violations of the trial protocol had been omitted from the original 2004 publication. Once these were accounted for, citalopram seemed no more effective than a placebo.

Both companies admitted that they had misrepresented safety and efficacy data, and paid heavy fines. Yet, Jureidini and McHenry point out, GlaxoSmithKline continued to claim that the findings of Study 329 had been accurately reported. And the FDA, they say, has taken no action to correct misreporting of Study CIT-MD-18 in Forest's application to license escitalopram to treat adolescent depression.

Companies hand over raw trial data only if forced, usually in the course of litigation (which they budget for). Despite attempts to make the process more transparent, for example by mandating the preregistration of clinical trials, many of those data are not in the public domain. That's why, the authors believe, these cases represent the tip of an iceberg.

Falsifiable theory

The authors agree that the randomized, placebo-controlled trial is the best method we have for testing drugs, and they argue that every scientific theory should be tested by, in Popper's phrase, attempting to falsify the null hypothesis. In a trial, this means trying to disprove the idea that the treatment makes no difference. Adhering to this principle, researchers can never say for sure that a treatment is effective, but they can say definitively that it is not effective.

However, the authors charge that drug companies have made even that impossible, by designing protocols that guarantee a positive outcome or by spinning a negative one. One concern is the redefinition of endpoints mid-trial — a worry that resurfaced in the context of the US National Institute of Allergy and Infectious Diseases' ongoing trial of the potential COVID-19 drug remdesivir, made by Gilead Sciences of Foster City, California. Partial solutions, such as requiring companies to deposit trial results in public databases, haven't worked. The commercial disincentives are just too strong.

Popper's ideas have often been criticized. Theories are never truly falsified, critics say, just shown to be less wrong than others. But we've gone too far down the road to relativism, counter Jureidini and McHenry; Popper offers a standard of integrity to which we must return. The only way to ensure that, they conclude, is to have trials conducted in a public-health system or by an independent institution funded by a tax on the industry. This would work only with government support, which has been lacking. Yet models do exist. The Mario Negri Institute for Pharmacological Research in Milan, Italy, has been conducting independent clinical trials for nearly 60 years.

The current pandemic might provide the perfect opportunity to acknowledge that there is a problem: ill people need treatments and the well need a vaccine. Quoting ancient Greek historian Thucydides, the authors write: "There will be justice … when those who are not injured are as outraged as those who are."

Laura Spinney is a science writer based in Paris. Her most recent book is *Pale Rider: The Spanish Flu of 1918 and How it Changed the World*.
e-mail: lfspinney@gmail.com

Soldiers involved in investigating the poisoning of Sergei and Yulia Skripal in Salisbury, UK, in 2018.

RUFUS COX/GETTY

Nerve agents: from discovery to deterrence

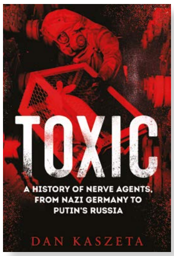
Chemical weapons treaties are not enough — scientists and industry play a part, too. **By Leiv K. Sydnes**

When the Russian former military officer Sergei Skripal and his daughter Yulia were poisoned with a 'novichok' nerve agent in the tranquil UK city of Salisbury in March 2018, it led to widespread fear that similar mysterious chemicals, illegal under international conventions, might be deployed elsewhere. What were they, where did they come from and what made them so deadly?

Enter *Toxic*, a round-up of the invention, production, proliferation and use of nerve agents. Author Dan Kaszeta has spent a career in defence and security, specializing in chemical, biological, radiological and nuclear materials. He worked for the US military, government and secret service before moving to the United Kingdom and becoming a security consultant. Drawing on this experience and an array of authoritative documents, he follows

the development of these deadly compounds, from the first discoveries in Germany in the 1930s to the Salisbury attack. He shows how fear of nerve agents influenced world events, such as the invasion of Iraq in 2003. And he reminds us that even with an international convention banning them, the threat of chemical weapons being used outside conventional warfare is ever-present. Stronger on the early history than on more recent politics, *Toxic* is a useful introduction to the subject.

The first nerve agents were the unintended consequence of a civil research project to make insecticides to secure Germany's food supply, led by Gerhard Schrader at chemical and pharmaceuticals giant IG Farben in Leverkusen from 1934. The research produced a liquid compound, eventually called tabun, with a toxicity far beyond anything seen before (0.1 milligrams of tabun per kilogram of body weight killed the monkeys used for testing). Almost at once, the Nazi authorities launched a massive programme, involving university scientists, the chemical industry and military personnel, to develop weapons that would distribute nerve agents – first tabun and later sarin – in a predictable manner in combat, without harming those who applied them.



Toxic: A History of Nerve Agents, From Nazi Germany to Putin's Russia
Dan Kaszeta
Hurst (2020)

> **"Manufacturers of the key ingredient refused to supply goods for use in weapons production."**

This happened even though chemical weapons were not part of German military strategy. Two factors were decisive. First was a capable industrial partner: IG Farben was available and willing. Second, the German leadership believed that the United States had nerve agents and could retaliate if attacked. I don't think these two factors are as independent as they seem. People such as IG Farben executive Otto Ambros, who kept Adolf Hitler informed about perceived US weapons capabilities, were also central in running the industrial production of the nerve agents, so had a personal business interest. Ambros made a fortune, even though the Nazi military did not, in the event, deploy the weapons.

Ethical standards

What would have happened had the German chemical industry declined to become involved? Kaszeta does not discuss this, but perhaps nerve-agent weapons would not have been developed. That was the case for the decades-long US programme to develop 'binary' chemical weapons, which contain precursor compounds that are mixed to produce the toxic agent on detonation. This programme was terminated around 1990, after Mobay and Occidental, the two domestic manufacturers of the key ingredient thionyl chloride, refused to supply it because company policy forbade the sale of goods for use in weapons production.

Unfortunately, the absence of such an ethical standard enabled Iraqi president Saddam Hussein to start large-scale production of tabun and sarin in 1981, using equipment and chemicals supplied by European and US companies.

After the first Gulf War (1990–91), Iraq's chemical weapons were destroyed, making the later Iraq War a significant moment in this shameful history. The invasion by the United States and its allies in March 2003 had the declared aim of eliminating weapons of mass destruction, in particular chemical weapons. None was found. In February that year, US secretary of state Colin Powell had presented 'evidence' for the presence of such weapons in Iraq before the United Nations security council; this is now generally accepted to have been based on faulty intelligence.

Kaszeta devotes just half a page to this conflict. He does not mention false or fabricated intelligence, and describes as "conventional wisdom" the statement that the invading forces did not discover any chemical weapons, leaving the impression that the weapons might indeed have existed. He seems to imply that Hussein's previous use of nerve agents justified his punishment. This was not an argument used by world leaders as they discussed an invasion. To me, the sacrifice of truth in the process of this discussion seemed shattering.

Toxic is at its best when the people making decisions, taking initiatives and violating agreements appear as characters in the chronological narrative, from the beginning to about 1950. At that point, Soviet and Western powers had acquired enough knowledge about nerve-agent production, through interrogation of German chemists and investigations of actual chemical weapons, to start large-scale production themselves.

Later sections are bogged down in detail. Yes, it is very difficult to turn tonnes of nerve agents into safe and effective weapons, but we don't need elaborate descriptions of artillery-shell testing to demonstrate the point. More important, bombs aren't necessary to cause panic and death. It was relatively easy for the Aum Shinrikyo cult to make and use sarin in Tokyo in 1995. Likewise, a skilled person with the right chemicals could prepare and apply novichoks. The 1997 Chemical Weapons Convention and its implementing body, the Organisation for the Prohibition of Chemical Weapons, cannot prevent this.

Kaszeta's book is informative; it should satisfy curious non-specialists able to digest details. The omission of chemical structures is generally no drawback, but when compounds are described as similar, some drawings in an appendix would have clarified the point. Yet the take-home message is clear: nerve agents are fairly easy to make, but difficult to turn into weapons. Their application in small quantities by aggressive individuals therefore remains a threat.

Leiv Sydnes is a professor of organic chemistry at the University of Bergen in Norway. He chaired the international task group that assessed the impact of scientific advances on the Chemical Weapons Convention in 2007 and 2012.
e-mail: leiv.sydnes@uib.no

Readers respond

Correspondence

Stop prevaricating, build in resilience

Investment in resilience is too frequently made only during or after a crisis – with the COVID-19 pandemic being one of the latest examples. A different approach is becoming ever more urgent if we are to secure the resilience of our society and natural resources (see, for example, *Nature* **581**, 119; 2020).

This approach must go beyond siloed strategies to include all five components of the system in which we live. These 'five capitals' are natural, human, social, built and financial, along with their interdependencies and feedbacks. They form a framework for sustainability, which will enable long-term planning for global resilience.

Such an approach would involve a shift from classifying the probability and consequence of known threats to addressing multiple hazards and recoverability. Emergent and interconnected issues, including adaptive capacity in organizations and critical infrastructure, must be actively managed. And we need to find ways to get company boards, governments and society in general to invest in resilience – even when there is not yet an economic argument for doing so (see also G. K. Marinov *Nature* **581**, 262; 2020).

Long-term planning and investment can be guided by short-term emergency responses, effective adaptation to repeated shocks and proper preparation for unexpected events (H. Weise *et al. Oikos* **129**, 445–456; 2020).

Jim A. Harris* Cranfield University, Bedford, UK.
j.a.harris@cranfield.ac.uk
*On behalf of 6 correspondents: see go.nature.com/2cawjjt


MARCIA BJORNERUD

Geoscientist's snap of *Nature* covers

As a long-time subscriber to *Nature*, with full access to the journal online as well as in print, I resolved to renounce my lingering allegiance to printed editions. While I was tipping old issues into the recycling bin, scores of vivid covers spanning decades of scientific advances caught my imagination. I decided to repurpose them into a striking collage (pictured).

As a geoscientist, I instinctively put the Earth at the centre, with the word 'nature' spiralling out of it in a potent incantation. Three eyes – of a baby squid, a human and a hurricane – indicate sentient life. Zooming in on the globe's surface reveals words that mark Earth's kaleidoscope of attributes (ice, clouds, bacteria, evolution and so on). The halo of radial and concentric colour gradients represents the interacting environmental gradients – rarely sharp boundaries – that define regions and ecosystems around the planet. Macroscopic and microscopic images are juxtaposed in the mosaic to reflect the complexity of natural systems at every scale.

And Charles Darwin floats, god-like, in the upper right, in an image that is itself a mosaic. It is taken from the cover of *Nature*'s 19 November 2009 issue (see go.nature.com/31vjv5t), which marked the 150th anniversary of *On the Origin of Species*.

Marcia Bjornerud Lawrence University, Appleton, Wisconsin, USA.
marcia.bjornerud@lawrence.edu

Legacy of a young Black professor

The story of our colleague, 35-year-old biology professor Lynika Strozier, is a sobering reminder of the hurdles faced by Black American scientists (*Nature* **582**, 147; 2020). Her death from COVID-19 complications during a global pandemic and nationwide reckoning of systemic racism has led to an outpouring of support.

Strozier was raised by her grandmother and was diagnosed with a learning disorder at an early age. She went on to earn two master's degrees simultaneously, from Chicago's Loyola University and University of Illinois. As she told the *Chicago Tribune* in 2012, "You get knocked down so many times, you learn to pick yourself up."

Strozier overcame these challenges through hard work, perseverance and strong relationships. She said that her research into biodiversity had endowed her with a previously unimaginable confidence. These experiences made her a keen mentor of other young researchers.

Her family started a GoFundMe campaign (https://gf.me/u/x737pr) to help offset Strozier's medical and funeral costs. Because this quickly surpassed expectations, they have now created a scholarship fund in her name.

Sushma Reddy University of Minnesota, St Paul, Minnesota, USA.

Ylanda Wilhite, **Matt Von Konrat** Field Museum of Natural History, Chicago, Illinois, USA.
ywilhite@fieldmuseum.org

Expert insight into current research

News & views

Quantum optics

Quantum fluctuations affect macroscopic objects

Valeria Sequino & Mateusz Bawaj

A method has been reported that improves the precision of measurements made by gravitational-wave detectors beyond an intrinsic limit — and shows that quantum fluctuations can alter the position of macroscopic objects. **See p.43**

In the hands of skilled experimentalists, light can be used as a probe for extremely precise measurements. However, the quantum nature of light places an intrinsic limit on the precision of such measurements. On page 43, Yu *et al.*[1] report that this limit has been overcome in experiments carried out using the Laser Interferometer Gravitational-Wave Observatory (LIGO) at Livingston, Louisiana. Moreover, the authors report the measurement of the effects of quantum fluctuations on macroscopic, kilogram-mass objects at room temperature. This is remarkable, because such fluctuations occur at size scales that are comparable to the dimensions of elementary particles.

Exceptionally sensitive detectors known as interferometers are used to measure the small distance variations induced by gravitational waves, which are produced by some of the most catastrophic events in the Universe. In the LIGO interferometer, mirrors are placed on kilogram-mass test objects at either end of two 4-kilometre-long cavities (arms); each pair of mirrors forms a system called an optical cavity. To attenuate external noise, the test masses are suspended on pendulums, which can oscillate only with frequencies that are much smaller than the frequency of the gravitational signal they are used to detect. Laser light is split into two beams, which are each sent down a different arm and reflected between the mirrors in the cavity. When the beams leave the cavity, they are recombined to produce interference patterns, which are then analysed for evidence of gravitational waves.

Light is electromagnetic radiation, and the lowest-energy quantum state of the electromagnetic light field is known as the vacuum. Despite its name, this vacuum is not completely empty. It contains quantum fluctuations that produce uncertainties in measurements of the amplitude and phase of light waves (in the case of a sinusoidal wave, the phase describes the shift of the waveform away from the minimum amplitude that corresponds to the start of the wave cycle). These uncertainties are quantified by Heisenberg's uncertainty principle.

Vacuum fluctuations cause noisy readouts in precision measurements made using light. Fluctuations in the measurements of the phase of light produce a phenomenon known as shot noise, whereas fluctuations in measurements of the amplitude of light produce radiation-pressure noise. The combination of these two is called quantum noise, and it limits the precision of measurements of tiny forces and displacements. The highest precision of any measurement that can be achieved using naturally occurring quantum states is called the standard quantum limit (SQL).

The SQL is a direct consequence of the Heisenberg uncertainty principle, which states that it is not possible to measure the position and momentum of an object simultaneously with unlimited precision. An electromagnetic field can be mathematically described as a set of two oscillating components: one component is related to the amplitude, and the other to the phase, of the wave. The fluctuations of these two also obey the Heisenberg uncertainty principle. However, the precision of measurements of amplitude and phase can be greatly improved if the magnitudes of uncertainties regarding the two components correlate with each other (Fig. 1). Such correlations arise spontaneously when light travels in suspended interferometers, such as the one

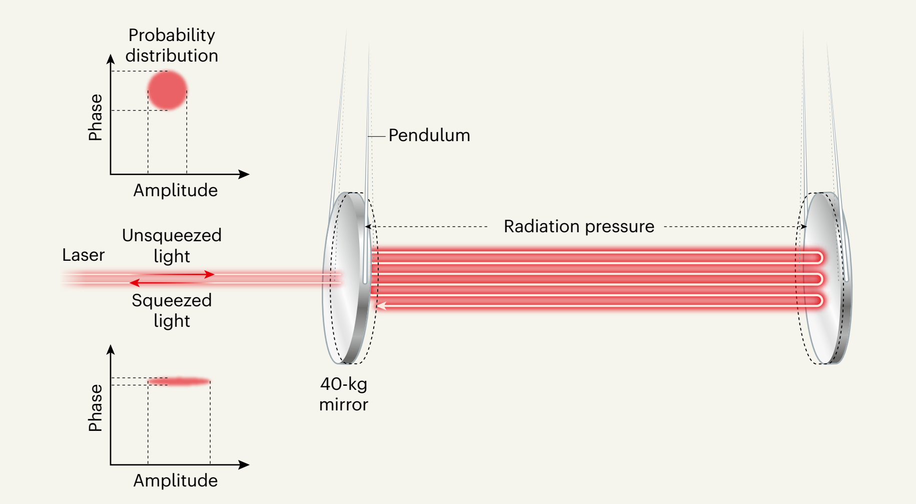


Figure 1 | Light squeezing due to the ponderomotive effect in one arm of a gravitational-wave detector. Gravitational-wave detectors contain optical cavities, which consist of mirrors suspended from pendulums and separated by distances of several kilometres. Light enters the cavity in an 'unsqueezed' state — that is, quantum fluctuations related to the phase and amplitude of light (uncertainties in the probability distribution of measurements) do not correlate with each other. The oscillating movement of the mirrors, induced by the radiation pressure of circulating light, causes a phase shift of light trapped in the cavity, and generates quantum correlations between the amplitude and phase (termed the ponderomotive effect). Light exiting the cavity is therefore squeezed; for this example, the phase uncertainty has been reduced, whereas the amplitude uncertainty has increased. At a different observation frequency of the signal, light might be squeezed another way — with increased phase uncertainty and decreased amplitude uncertainty. Yu *et al.*[1] show that this effect can be used to increase the precision of measurements made by a gravitational-wave detector, thereby surpassing an intrinsic limit on precision (the standard quantum limit). The authors also show that radiation pressure noise — the minuscule variation of the force exerted on the kilogram-scale mirrors by light trapped in the cavity — contributes to the motion of the suspended mirrors.

used by LIGO. Suspended interferometers measure the phase of the output field of light waves, which is affected by both amplitude and phase fluctuations of the input vacuum field. This correlation is called the ponderomotive effect[2]. The detection response of the instrument is frequency dependent, and the effects of the amplitude fluctuations are more evident in the low-frequency realm of the detection band, whereas the phase fluctuations are more evident at high frequencies.

Light that has correlations between the uncertainties of its amplitude and phase is said to be 'squeezed'. The Heisenberg principle still holds for squeezed light states, but when one of the uncertainties is reduced, the other is increased. Squeezed light can be used in experiments to reduce the uncertainty of one of the correlated parameters. A special case of squeezed light, known as the squeezed vacuum, forms when the average amplitude of the light is zero.

Phase-squeezed light, in which the uncertainty associated with the phase is squeezed, has been used to reduce shot noise for both LIGO[3] and Virgo, the gravitational-wave detector located in Cascina, Italy[4]. And the ponderomotive effect has previously been demonstrated using the mechanical motion of pico- to microgram-scale mirrors in laboratory experiments[5,6]. Yu *et al.* now confirm that the ponderomotive effect occurs in the optical cavities of the LIGO interferometer, and have investigated whether it can be used in combination with squeezed-vacuum states to reduce quantum noise below the SQL in measurements of mirror position in the cavities.

The authors measured the noise in the LIGO interferometer under two sets of experimental conditions: one in which squeezed-vacuum states were injected into the output port of the interferometer, and another in which squeezed-vacuum states were not injected. They then plotted sensitivity curves for the data, which chart the noise level in the detector and define the minimum gravitational signal that can be detected as a function of the signal's frequency. This revealed that, once classical (non-quantum) noise had been subtracted from their data, the uncertainties in the phases of the laser beam and in the positions of the mirrors produce a combined quantum noise below the SQL. Yu and colleagues have therefore demonstrated two fundamental points: that quantum fluctuations of light exert a measurable force on macroscopic objects (the 40-kg mirrors); and that the quantum noise corresponding to these disturbances can be reduced to below the SQL.

One of the main difficulties for these kinds of measurement is thermal fluctuations — which can drive mirror motion and are one of the main sources of noise for gravitational-wave detectors. Cryogenic conditions have therefore been needed in some previously reported experiments[7,8] to reduce quantum noise to less than the SQL. Impressively, Yu and co-workers' measurements were made at room temperature.

Yu *et al.* are the first to have proved experimentally that a quantum non-demolition technique — a method in which a measurement of a quantum system is performed repeatedly without perturbing it[9] — works in gravitational-wave detectors. At present, such detectors use phase-squeezed vacuum states to reduce shot noise, without considering the correlations that are introduced by the interferometer mirrors. This approach improves sensitivity only for gravitational signals in which the frequency is higher than 100 hertz, up to the limit of the detection band[6]. By contrast, Yu and colleagues' technique potentially enables broadband detection improvement. However, further work will be needed to reduce the classical noise in the interferometer.

Once better sensitivity has been developed, more gravitational waves could be detected than is possible at present. Future work in noise suppression will therefore take us towards an exciting era of sub-SQL performance of gravitational-wave detectors.

Valeria Sequino is in the Department of Physics, University of Naples Federico II, Naples 80126, Italy, and at INFN, Naples. **Mateusz Bawaj** is in the Department of Physics and Geology, University of Perugia, Perugia 06123, Italy, and at INFN, Perugia.
e-mail: sequino@na.infn.it; bawaj@pg.infn.it

1. Yu, H. *et al. Nature* **583**, 43–47 (2020).
2. Corbitt, T. *et al. Phys. Rev. A* **73**, 023801 (2006).
3. Tse, M. *et al. Phys. Rev. Lett.* **123**, 231107 (2019).
4. Acernese, F. *et al. Phys. Rev. Lett.* **123**, 231108 (2019).
5. Kronwald, A., Marquardt, F. & Clerk, A. A. *N. J. Phys.* **16**, 063058 (2014).
6. Di Pace, S. *et al. Proc. Gravitational-wave Sci. Technol. Symp. 2019*. Available at http://dx.doi.org/10.5281/zenodo.3569196 (2020).
7. Møller, C. B. *et al. Nature* **547**, 191–195 (2017).
8. Teufel, J. D., Lecocq, F. & Simmonds, R. W. *Phys. Rev. Lett.* **116**, 013602 (2016).
9. Braginsky, V. B., Vorontsov, Y. I. & Thorne, K. P. *Science* **209**, 547–557 (1980).

Tumour biology

A 'safety net' causes cancer cells to migrate and grow

Emma Nolan & Ilaria Malanchi

Immune cells called neutrophils can support the spread of cancer. How neutrophils aid this process now comes into focus through insights into the function of structures called neutrophil extracellular traps. **See p.133**

A neutrophil is a type of immune cell that provides the body with one of its first lines of defence against infection. However, in many contexts, neutrophils also have the ability to promote metastasis — the migration of cancer cells from their primary site and their growth in other locations in the body. On page 133, Yang *et al.*[1] shed light on how neutrophils aid this deadly process.

A key feature of neutrophils is their ability to extrude a structure called a neutrophil extracellular trap (NET) into their surroundings (Fig. 1). This consists of a web of DNA coated in enzymes toxic to microorganisms, and it can trap and kill invading microbes. But in the lungs, NETs are induced by inflammation, and their tumour-boosting activity has been linked to NET-associated enzymes[2]. A growing body of evidence indicates that NETs mediate the development and enhancement of the invasive properties of cancer cells[3], but how they boost metastasis has remained largely unknown. Moreover, a mechanism that enables cancer cells to sense NETs has not been reported previously. Yang *et al.* now provide much-needed insight into the tumour-promoting effects of these traps.

The authors began by assessing NETs in primary and metastatic tumours from 544 people with breast cancer. NETs were scarce at primary-tumour sites, but were abundant in the liver — a common site of breast cancer spread. Importantly, the authors found an association between higher levels of NET DNA in the blood of people with early-stage breast cancer and subsequent metastasis of the cancer to the liver. This indicates that monitoring NET DNA in blood samples might be a way of assessing disease prognosis.

To investigate the relationship between NETs and cancer cells *in vivo*, the authors transplanted breast cancer cells of human or mouse origin into mice, and analysed metastatic tumour cells. They found that NETs accumulated in the liver in both mouse models tested. The finding is consistent with the results of the authors' analysis of tumours from people with cancer.

Yang *et al.* report that, in their mouse models, NETs were induced in the liver

https://doi.org/10.1038/d41586-020-01672-3

News & views

Tumour biology

A 'safety net' causes cancer cells to migrate and grow

Emma Nolan & Ilaria Malanchi

Immune cells called neutrophils can support the spread of cancer. How neutrophils aid this process now comes into focus through insights into the function of structures called neutrophil extracellular traps.

A neutrophil is a type of immune cell that provides the body with one of its first lines of defence against infection. However, in many contexts, neutrophils also have the ability to promote metastasis – the migration of cancer cells from their primary site and their growth in other locations in the body. Writing in *Nature*, Yang *et al.*[1] shed light on how neutrophils aid this deadly process.

A key feature of neutrophils is their ability to extrude a structure called a neutrophil extracellular trap (NET) into their surroundings (Fig. 1). This consists of a web of DNA coated in enzymes toxic to microorganisms, and it can trap and kill invading microbes. But in the lungs, NETs are induced by inflammation, and their tumour-boosting activity has been linked to NET-associated enzymes[2]. A growing body of evidence indicates that NETs mediate the development and enhancement of the invasive properties of cancer cells[3], but how they boost metastasis has remained largely unknown. Moreover, a mechanism that enables cancer cells to sense NETs has not been reported previously. Yang *et al.* now provide much-needed insight into the tumour-promoting effects of these traps.

The authors began by assessing NETs in primary and metastatic tumours from 544 people with breast cancer. NETs were scarce at primary-tumour sites, but were abundant in the liver – a common site of breast cancer spread. Importantly, the authors found an association between higher levels of NET DNA in the blood of people with early-stage breast cancer and subsequent metastasis of the cancer to the liver. This indicates that monitoring NET DNA in blood samples might be a way of assessing disease prognosis.

To investigate the relationship between NETs and cancer cells *in vivo*, the authors transplanted breast cancer cells of human or mouse origin into mice, and analysed metastatic tumour cells. They found that NETs accumulated in the liver in both mouse models tested. The finding is consistent with the results of the authors' analysis of tumours from people with cancer.

Yang *et al.* report that, in their mouse models, NETs were induced in the liver before metastatic cells could be detected there. The authors show that the efficiency with which cancer cells metastasized to the liver depended on NETs, because metastasis in mice was substantially impaired on removal of NETs, either by means of the DNA-degrading enzyme DNase I or if the animals were genetically engineered to lack an enzyme required for NET formation[4].

Previous work[5] has led to the proposal that NET-dependent metastasis to the liver occurs through an indirect mechanism by the physical trapping of 'passer-by' cancer cells by NETs. Yang and colleagues showed that NET DNA directly stimulated the migration and adhesion of human breast cancer cells when tested *in vitro*.

The authors next sought to discover how this migratory behaviour is induced. By adding a tag to NET DNA and using it as bait with which to capture and identify proteins with which it interacts, they found a receptor called CCDC25 that could bind to NET DNA. It is present on the surface of cancer cells, and Yang and colleagues report that CCDC25 could bind to NET DNA with high specificity and affinity, enabling 'NET sensing' by cancer cells. Impressively, the authors identified the specific extracellular portion of CCDC25 that binds to NET DNA.

The authors confirmed that NET-mediated stimulation of cancer-cell migration is driven by CCDC25 by showing that depleting it

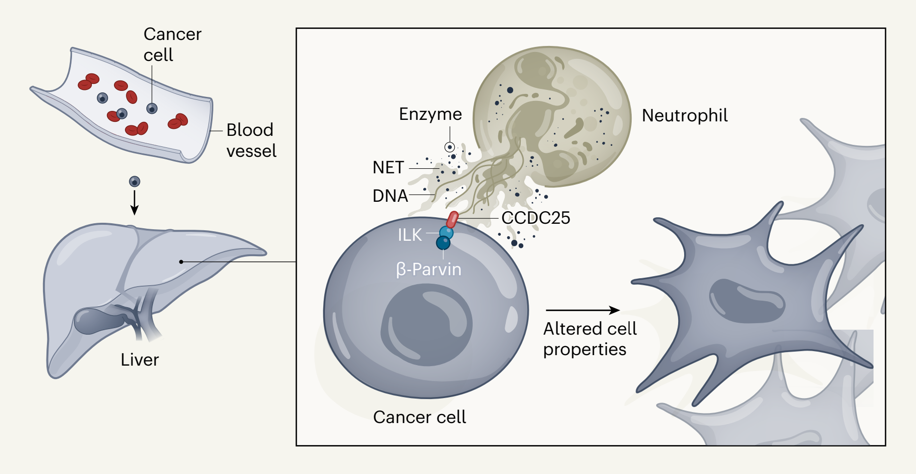


Figure 1 | A process that aids the spread of cancer cells. Cancer-cell migration (metastasis) from the primary site of growth, through the bloodstream, to a secondary site requires supportive signals in that distant organ. From their study of mice and samples from people with cancer, Yang *et al.*[1] identify a mechanism that enables cancer cells to metastasize to the liver. Before metastatic cancer cells arrive in the liver, immune cells there called neutrophils extrude a structure called a neutrophil extracellular trap (NET), which contains DNA and enzymes that kill microorganisms. Yang and colleagues report that this structure binds to a protein called CCDC25 on the cancer-cell surface. This interaction triggers a signalling cascade in the tumour cell that is mediated by the enzyme ILK and its partner, the protein β-parvin. This pathway modifies characteristics of the cancer cell and alters actin filaments in the cytoplasm (not shown) that affect cell shape. The changes that occur increase the cancer cell's adhesive and invasive properties and boost its proliferation.

from human breast cancer cells cultured *in vitro*, or from samples of patients' primary breast-tumour cells, drastically reduced migration of the cancer cells when tested *in vitro*. Compared with the case for mice in which CCDC25 was still present, eliminating CCDC25 from the surface of cancer cells in mice significantly lessened the development of metastasis to the liver and decreased metastasis to the lungs after inflammation-inducing treatment with the molecule lipopolysaccharide (LPS). The role of LPS in triggering lung metastasis associated with NETs was previously reported[2]. Yang *et al*. observed similar reductions in metastasis to the lung on LPS treatment in their experiments if they used animals obtained by crossing mice lacking CCDC25 with mice that model spontaneously forming breast cancer, called MMTV-PyMT mice. Interestingly, the role of the interaction between CCDC25 and NET DNA in supporting metastasis in the lungs might occur only in the context of infection, whereas its effect on liver metastasis might occur spontaneously.

Finally, Yang and colleagues reveal how tumour cells profit from this interaction with NETs. Using CCDC25 as 'bait' in a biochemical technique to fish out CCDC25-interacting proteins in cancer cells, they identified one such protein — integrin-linked kinase (ILK), an enzyme that regulates processes such as cellular migration and proliferation[6]. When ILK was removed or its downstream signalling partner, the protein β-parvin, was disabled, cancer-cell growth and motility *in vitro* were substantially impaired and, in mice, metastasis to the liver was reduced. Together, the authors' results indicate that the binding of NET DNA to CCDC25 enhances aggressive cancer-cell behaviour by activating an ILK-mediated signalling cascade.

Yang *et al*. show that the ability of NET DNA to foster metastasis to the liver was not specific to breast cancer cells. NETs were observed in liver metastases in people with colon cancer and in tumours arising in the livers of mice that had been injected with human colon cancer cells. The authors found that if human breast and colon cancer cells were engineered to increase their levels of CCDC25, this helped to fuel liver metastasis in mice given such cells. Crucially, the authors identified a correlation between high CCDC25 abundance in primary tumours and shorter long-term survival in patients across multiple cancer types, indicating that monitoring CCDC25 expression might be useful for predictive purposes.

Future studies will be needed to assess the feasibility of targeting CCDC25 for anticancer therapy. The expression of CCDC25 in different cell types and its possible functions in normal cells should be examined. Given that the authors have identified the precise extracellular portion of CCDC25 that interacts with NET DNA, it might be possible to develop specific inhibitors to block this interaction. Such a targeted approach would have the advantage of preserving other functions of NETs that help to fight infections.

It remains to be determined why the liver is particularly prone to NET accumulation compared with other metastatic sites. In the context of mammalian intestinal cancer, the release of NETs from neutrophils is linked to upregulation of a protein called complement C3a, which is mainly produced in the liver, and which can bind to a receptor on neutrophils[7]. Activation of the complement pathway occurs in the mammalian liver before the development of liver metastasis[8], and so a complement-dependent stimulation of NET formation could be hypothesized. However, the specific mechanism involved remains to be elucidated.

Yang and colleagues' findings represent a key advance in efforts to curb cancer spread, and might lead to the development of a specific strategy to halt NET boosting of cancer metastasis. Moreover, the data presented point to a possible way to predict metastasis to the liver by monitoring NET DNA in the blood.

Emma Nolan and **Ilaria Malanchi** are at the Francis Crick Institute, London NW1 1AT, UK.
e-mail: ilaria.malanchi@crick.ac.uk

1. Yang, L. *et al. Nature* https://doi.org/10.1038/s41586-020-2394-6 (2020).
2. Albrengues, J. *et al. Science* **361**, eaao4227 (2018).
3. Park, J. *et al. Sci. Transl. Med.* **8**, 361ra138 (2016).
4. Li, P. *et al. J. Exp. Med.* **207**, 1853–1862 (2010).
5. Cools-Lartigue, J. *et al. J. Clin. Invest.* **123**, 3446–3458 (2013).
6. Zheng, C.-C. *et al. Am. J. Cancer Res.* **9**, 186–197 (2019).
7. Guglietta, S. *et al. Nature Commun.* **7**, 11037 (2016).
8. Reis, E. S., Mastellos, D. C., Ricklin, D., Mantovani, A. & Lambris, J. D. *Nature Rev. Immunol.* **18**, 5–18 (2018).

https://doi.org/10.1038/d41586-020-01600-5

News & views

Neuroscience

Flipping the switch on the thermoregulatory system

Clifford B. Saper & Natalia L. S. Machado

A population of excitatory neurons has been found to have a key role in controlling body temperature in rodents. The discovery adds to a body of work that is raising questions about long-standing models of thermoregulation.

Body temperature in mammals is tightly controlled[1], and is typically maintained to within about 0.5° C of an animal's mean core temperature (usually around 37° C) throughout life. However, lack of food can cause some mammals to enter a sleep-like state – called torpor or hibernation, depending on its duration – in which the body temperature can drop by 5–10° C (or, in some cases, even more) to conserve energy. Animals can also increase their body temperature (fever) in response to infection, slowing the replication of some invaders and improving the animal's chance of survival[2]. It has long been known that these regulatory feats are accomplished by neurons in the brain's preoptic area, but the exact identities of those neurons and their connections have not been understood. Writing in *Nature*, Takahashi *et al.*[3] and Hrvatin *et al.*[4] add to a flurry of papers that are revolutionizing our understanding of the preoptic neurons at the heart of thermoregulation.

Decades of research into thermoregulation have produced a model[1] whereby excitatory neurons in a region of the preoptic area called the median preoptic nucleus are activated by warming of the skin. In the model, these excitatory neurons activate inhibitory neurons in an adjacent brain region, the medial preoptic area. The inhibitory neurons express an enzyme called glutamic acid decarboxylase (GAD), which synthesizes the inhibitory neurotransmitter molecule GABA. The model posits that these neurons, which are presumably GABA-releasing (GABAergic), then project to other regions of the brain, where they inhibit neurons that promote heat production and conservation. Thus, activity of the GABAergic neurons results in cooling (Fig. 1a). In cool ambient temperatures, the preoptic GABAergic neurons would be inhibited, releasing constraints on heat production and conservation. And during an inflammatory illness, the lipid prostaglandin E2 is released and acts on EP3 receptor proteins in the neurons, causing their inhibition and so promoting responses involving fever[5,6].

However, this model began to show cracks after a series of studies demonstrated that the preoptic neurons that cause cooling when activated express specific genetic markers, including several that encode certain protein fragments[7] and receptors[6,8,9]. Analysis of these markers revealed that the key neurons causing hypothermia are located in the median preoptic nucleus, but not in the medial preoptic area. These studies also found that the hypothermia-causing preoptic neurons release the excitatory molecule glutamate as their main neurotransmitter, rather than GABA. Like many other cells in the preoptic area, these neurons contain GAD, but they don't express the vesicular GABA transporter (Vgat) protein, which is needed to load GABA into synaptic vesicles, enabling the neurotransmitter to be released from the cell. Instead, the neurons express the vesicular glutamate transporter 2 (Vglut2) protein, making them glutamate-releasing (glutamatergic) and excitatory, rather than inhibitory[10].

Against this background, Takahashi and colleagues describe a genetic marker in mice for a particularly potent subset of preoptic thermoregulatory neurons: a gene that encodes the protein fragment pyroglutamylated RF-amide peptide (QRFP). The authors genetically engineered mice such that neurons that express QRFP could be activated by the injection of a small molecule, clozapine *N*-oxide. This chemogenetic activation caused the animals to become immobile, and led to a fall in body temperature to about 23–24° C. This hypothermia was associated with slowed heart rate and respiration, as well as reduced metabolic rate, similar to that seen in torpor or hibernation.

The authors next produced mice in which specific QRFP neurons or their projections (axon terminals) could be activated by laser light. A torpor-like state was produced when the authors used this optogenetic activation to

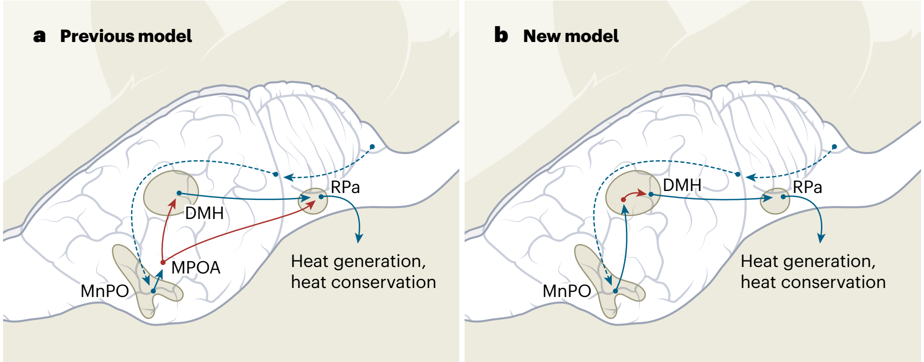


Figure 1 | Pathways for thermoregulation. Thermoregulation involves complex networks of excitatory and inhibitory neurons (blue and red dots, respectively, with their projections indicated by arrows). **a**, An existing model posits that warming of the skin leads (through neuronal pathways indicated by dashed arrows) to activation of excitatory neurons in a brain region called the median preoptic nucleus (MnPO). These neurons activate inhibitory neurons in the adjacent medial preoptic area (MPOA), which project to two more brain regions – the dorsomedial hypothalamus (DMH) and the raphe pallidus (RPa) – that generate and conserve heat. Thus, skin warming would cause a regulated reduction in body temperature. **b**, Takahashi *et al.*[3] and Hrvatin *et al.*[4] build on a flurry of work that supports an alternative model. The groups find that a population of mostly excitatory neurons in the MnPO causes profound hypothermia. Takahashi *et al.* show that these neurons express the protein fragment QRFP (not shown). The neurons connect directly to the DMH, where they presumably contact local inhibitory neurons to exert their effect on thermoregulation.

stimulate either the QRFP cell bodies or their terminals in the dorsomedial hypothalamus. This brain area is known to send projections to a site in the brain's medulla, the raphe pallidus, which promotes increases in body temperature. Interestingly, the body temperature of animals that had QRFP-neuron-induced hypothermia was still regulated, but around a lower setpoint. This situation is typical of hibernation, and suggests that, when the QRFP neurons are actively lowering body temperature, other non-QRFP preoptic neurons could continue to regulate body temperature, but at a lower level.

Takahashi *et al.* showed that nearly 80% of the QRFP neurons expressed Vglut2 but not Vgat. By contrast, only about 7% expressed Vgat but not Vglut2, and around 13% expressed both. Deleting Vgat from the QRFP neurons slightly slowed the initial fall in body temperature caused by activating these cells, but body temperature reached a level comparable to that of control animals after six hours. Deleting Vglut2, by contrast, prevented a hibernation-like state. Thus, the hypothermia produced by QRFP neurons is predominantly mediated by glutamatergic transmission.

Previous research has indicated that many of the preoptic neurons that drive hypothermia express proteins called pituitary adenylate cyclase-activating peptide (PACAP) and brain-derived neurotrophic factor (BDNF)[7]. Takahashi and co-workers demonstrated that most preoptic QRFP-expressing neurons also expressed BDNF and PACAP. However, about 75% of the BDNF and PACAP neurons in the median preoptic nucleus did not express QRFP. Similarly, 75% of the QRFP neurons expressed the EP3 receptor, but many EP3-expressing neurons did not express QRFP.

Hrvatin and co-workers used a different approach. The group analysed a marker of neuronal activity to determine the neuronal populations activated during torpor caused by 24 hours of food deprivation. The active neurons were distributed similarly to QRFP cells, and many of them expressed PACAP. Thus, the results of the two studies, taking very different approaches, reinforce one another.

Taken together, these observations suggest that there are several subpopulations of thermoregulatory neurons clustered together in the median preoptic nucleus, each distinguished by a unique pattern of gene expression. Among these, the QRFP group seems to be particularly important for producing deep hypothermia. This process is necessary when animals do not have sufficient food available to maintain their typical levels of metabolism and activity. At such times, animals can undergo daily torpor (brief periods when their body temperature might drop to 30° C or lower for a few hours, frequently seen in mice and rats) or hibernation (long, seasonal periods of deeper hypothermia, such as is seen in bears).

If similar groups of QRFP-expressing neurons are found in humans, they could represent a way to induce therapeutic hypothermia — for example, after heart attack or stroke, slowing down metabolic processes to help limit tissue damage. By contrast, during inflammatory illness, inhibition of QRFP glutamatergic neurons by the EP3 receptor might play a key part in producing fever[6]. Learning how to control these QRFP neurons could provide insight that will aid the development of new fever-reducing drugs.

The enormous range of body temperatures regulated by the QRFP neurons suggests that they and the other subpopulations of thermoregulatory neurons in the median preoptic nucleus might be the centrepiece of the brain's thermoregulatory system. But if these neurons are excitatory, and if they act on neurons that cause heat generation and conservation, then there must be an inhibitory link, almost certainly consisting of local inhibitory neurons called interneurons (Fig. 1b). This model calls for reconsideration of much of what we thought we knew about thermoregulation, in particular the physiological roles of genetically distinct subpopulations of median preoptic thermoregulatory neurons.

Clifford B. Saper and **Natalia L. S. Machado** are in the Department of Neurology, Division of Sleep Medicine, and Program in Neuroscience, Harvard Medical School, and Beth Israel Deaconess Medical Center, Boston, Massachusetts 02215, USA.
e-mail: csaper@bidmc.harvard.edu

1. Morrison, S. F. & Nakamura, K. *Annu. Rev. Physiol.* **81**, 285–308 (2019).
2. Blomqvist, A. & Engblom, D. *Neuroscientist* **24**, 381–399 (2018).
3. Takahashi, T. M. *et al. Nature* https://doi.org/10.1038/s41586-020-2163-6 (2020).
4. Hvratin, S. *et al. Nature* https://doi.org/10.1038/s41586-020-2387-5 (2020).
5. Lazarus, M. *et al. Nature Neurosci.* **10**, 1131–1133 (2007).
6. Machado, N. L. S., Bandaru, S. S., Abbott, S. B. G. & Saper, C. B. *J. Neurosci.* **40**, 2573–2588 (2020).
7. Tan, C. L. *et al. Cell* **167**, 47–59 (2016).
8. Yu, S. *et al. J. Neurosci.* **36**, 5034–5046 (2016).
9. Wang, T. A. *et al. Neuron* **103**, 309–322 (2019).
10. Moffitt, J. R. *et al. Science* **362**, eaau5324 (2018).

neurons are actively lowering body temperature, other non-QRFP preoptic neurons could continue to regulate body temperature, but at a lower level.

Takahashi *et al.* showed that nearly 80% of the QRFP neurons expressed Vglut2 but not Vgat. By contrast, only about 7% expressed Vgat but not Vglut2, and around 13% expressed both. Deleting Vgat from the QRFP neurons slightly slowed the initial fall in body temperature caused by activating these cells, but body temperature reached a level comparable to that of control animals after six hours. Deleting Vglut2, by contrast, prevented a hibernation-like state. Thus, the hypothermia produced by QRFP neurons is predominantly mediated by glutamatergic transmission.

Previous research has indicated that many of the preoptic neurons that drive hypothermia express proteins called pituitary adenylate cyclase-activating peptide (PACAP) and brain-derived neurotrophic factor (BDNF)[7]. Takahashi and co-workers demonstrated that most preoptic QRFP-expressing neurons also expressed BDNF and PACAP. However, about 75% of the BDNF and PACAP neurons in the median preoptic nucleus did not express QRFP. Similarly, 75% of the QRFP neurons expressed the EP3 receptor, but many EP3-expressing neurons did not express QRFP.

Hrvatin and co-workers used a different approach. The group analysed a marker of neuronal activity to determine the neuronal populations activated during torpor caused by 24 hours of food deprivation. The active neurons were distributed similarly to QRFP cells, and many of them expressed PACAP. Thus, the results of the two studies, taking very different approaches, reinforce one another.

Taken together, these observations suggest that there are several subpopulations of thermoregulatory neurons clustered together in the median preoptic nucleus, each distinguished by a unique pattern of gene expression. Among these, the QRFP group seems to be particularly important for producing deep hypothermia. This process is necessary when animals do not have sufficient food available to maintain their typical levels of metabolism and activity. At such times, animals can undergo daily torpor (brief periods when their body temperature might drop to 30° C or lower for a few hours, frequently seen in mice and rats) or hibernation (long, seasonal periods of deeper hypothermia, such as is seen in bears).

If similar groups of QRFP-expressing neurons are found in humans, they could represent a way to induce therapeutic hypothermia — for example, after heart attack or stroke, slowing down metabolic processes to help limit tissue damage. By contrast, during inflammatory illness, inhibition of QRFP glutamatergic neurons by the EP3 receptor might play a key part in producing fever[6]. Learning how to control these QRFP neurons could provide insight that will aid the development of new fever-reducing drugs.

The enormous range of body temperatures regulated by the QRFP neurons suggests that they and the other subpopulations of thermoregulatory neurons in the median preoptic nucleus might be the centrepiece of the brain's thermoregulatory system. But if these neurons are excitatory, and if they act on neurons that cause heat generation and conservation, then there must be an inhibitory link, almost certainly consisting of local inhibitory neurons called interneurons (Fig. 1b). This model calls for reconsideration of much of what we thought we knew about thermoregulation, in particular the physiological roles of genetically distinct subpopulations of median preoptic thermoregulatory neurons.

Clifford B. Saper and **Natalia L. S. Machado** are in the Department of Neurology, Division of Sleep Medicine, and Program in Neuroscience, Harvard Medical School, and Beth Israel Deaconess Medical Center, Boston, Massachusetts 02215, USA.
e-mail: csaper@bidmc.harvard.edu

1. Morrison, S. F. & Nakamura, K. *Annu. Rev. Physiol.* **81**, 285–308 (2019).
2. Blomqvist, A. & Engblom, D. *Neuroscientist* **24**, 381–399 (2018).
3. Takahashi, T. M. *et al. Nature* **583**, 109–114 (2020).
4. Hvratin, S. *et al. Nature* **583**, 115–121 (2020).
5. Lazarus, M. *et al. Nature Neurosci.* **10**, 1131–1133 (2007).
6. Machado, N. L. S., Bandaru, S. S., Abbott, S. B. G. & Saper, C. B. *J. Neurosci.* **40**, 2573–2588 (2020).
7. Tan, C. L. *et al. Cell* **167**, 47–59 (2016).
8. Yu, S. *et al. J. Neurosci.* **36**, 5034–5046 (2016).
9. Wang, T. A. *et al. Neuron* **103**, 309–322 (2019).
10. Moffitt, J. R. *et al. Science* **362**, eaau5324 (2018).

Condensed-matter physics

Atomic forces mapped out by lasers

Michael A. Sentef

The forces between electrons and nuclei in solids are difficult to image directly. A study shows that these forces can instead be indirectly imaged using the light emitted when the electrons are subjected to a strong laser field. **See p.55**

One of the central goals of physics is to gain a detailed understanding of nature's building blocks and the mutual forces between them. In materials, such building blocks are atomic nuclei and the electrons that zip around between these nuclei, with forces acting on atomic length scales. Direct imaging of such forces using light is notoriously difficult and typically requires X-ray wavelengths. However, on page 55, Lakhotia *et al.*[1] demonstrate that indirect imaging is possible using visible light, even though the wavelengths of this light are about 10,000 times larger than atomic scales.

The authors achieved this feat using a method called high-harmonic generation, in which a strong laser field provides the electrons with more energy than they need to overcome the forces pulling them back to the nuclei. The shaken electrons then emit light at multiples of the laser frequency, known as high harmonics. This emission is a consequence of the nonlinearity of the energy 'landscape' that the electrons are subjected to inside the periodic lattice of nuclei when they are driven by an intense laser field.

To understand this effect, consider playing a note on a trumpet. When the instrument is played at normal strength, a pure tone is heard at the intended frequency. However, when one blows the trumpet strongly, higher overtones emerge because the amplitude of the instrument's excitation is sufficiently large to probe a nonlinear regime.

Electrons in solids are quantum-mechanical objects described by a wavefunction that determines the probability of finding them at a specific position and with a particular velocity or momentum. For free particles, momentum is the product of mass and velocity. However, electrons in solids are not free, but are affected by the potential energy provided by the uniform atomic lattice. The electrical forces applied to the electrons by the nuclei are given by the slope of the potential-energy landscape at each position (Fig. 1a) and are analogous to the gravitational forces pulling back a hiker in the mountains. But how can these forces be mapped out by shaking the electrons with a laser?

The answer to this question is best understood by considering how an electron's energy depends on its momentum (Fig. 1b). The kinetic energy of a free electron grows quadratically with its velocity or momentum,

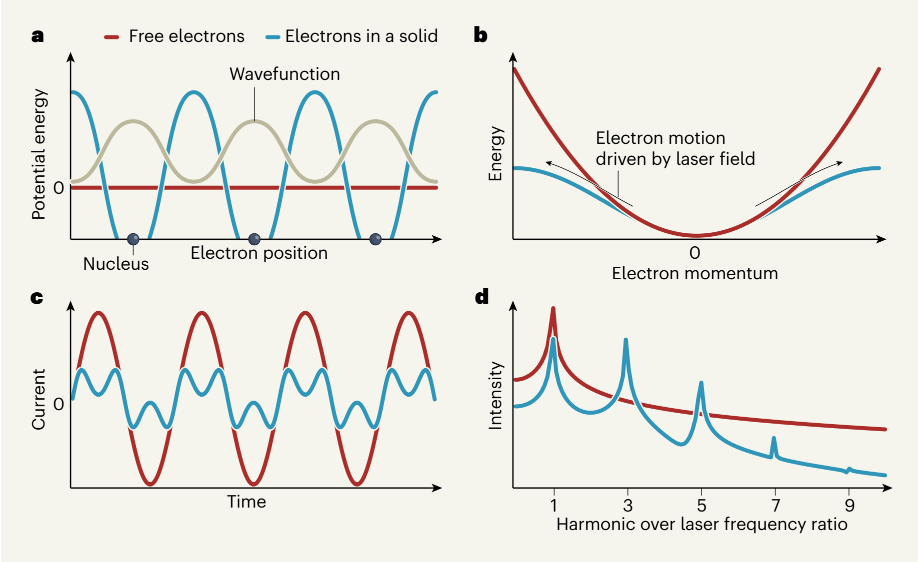


Figure 1 | High-harmonic generation. **a**, The potential energy of free electrons is zero, but that of electrons in a solid varies because these particles are attracted to nuclei located at the potential-energy minima. The wavefunction of such electrons has a periodicity determined by the positions of the nuclei. **b**, The energy–momentum relation for free electrons has the shape of a parabola. However, for electrons in a solid, the potential-energy 'landscape' changes this parabola into a shape that can be described as an energy band. When a strong laser field is applied to these band electrons, they are driven into the region of non-parabolicity. **c**, In such a field, the current of free electrons has sinusoidal oscillations, whereas that of band electrons shows deviations from these oscillations. **d**, The free electrons produce light at the laser frequency (the single peak present). Lakhotia *et al.*[1] show that the band electrons also emit light at odd multiples (high harmonics) of this frequency.

resulting in a curve known as a parabola. For an electron in a solid, the potential-energy landscape changes this parabola into an energy band that resembles the parabola at small momenta but flattens out when the electronic wavefunction reaches a momentum comparable to the inverse of the interatomic distance in the lattice. Such flattening of the energy–momentum curve corresponds to the nonlinearity that makes a trumpet play overtones.

To reach this nonlinear regime, one needs to apply a strong laser field that accelerates the electrons to large-enough momenta. Within the parabolic part of the energy band, the magnitude of the current produced by the electrons follows sinusoidal oscillations in the amplitude of the applied laser field in lockstep (Fig. 1c). However, once the nonlinearity is reached, the current deviates from sinusoidal behaviour and overtones start to emerge.

A simple way to see the connection between the non-parabolic part of the energy band and the emergence of overtones in the current is by noting that the velocity of the electrons is given by the slope of the energy–momentum curve. When the electrons are accelerated to high momenta, the band flattens out, the velocity decreases and the magnitude of the current is reduced. Because the band flattening is directly linked to the potential energy caused by forces between electrons and nuclei, the deviations from a sinusoidal current encode information about the energy landscape itself.

Lakhotia and colleagues' main achievement is the precise measurement of these deviations and the reconstruction of the underlying potential-energy landscapes inside the materials they considered. In practice, they did not record the electronic currents directly; rather, they measured the spectra of light emitted by the moving charges (Fig. 1d). These spectra contain a single peak at the laser frequency and additional peaks at odd high harmonics. The authors analysed in detail the heights of these peaks and the phases of the emitted light — the phase of a light wave specifies in which stage of an oscillation cycle the electric field of the wave is.

"The shaken electrons then emit light at multiples of the laser frequency, known as high harmonics."

To reconstruct the energy landscapes, Lakhotia *et al.* needed to assume that the atomic forces were weak compared with the driving force provided by the laser field[2]. This assumption seems to be fulfilled for the materials considered, partly because the atomic forces are not too strong. As a result, the deviation between the free-electron parabola and the flattened band is relatively small. An intriguing open question is whether a method known as high-harmonic spectroscopy[3] can be generalized to reveal detailed information about the forces inside solids when these forces are strong.

The authors also needed to assume the validity of the independent-electron picture, in which the mutual repulsion between electrons can be neglected. This picture is inappropriate for some materials more exotic than those studied here. For instance, in strongly correlated electronic materials, electron–electron interactions can lead to astonishing effects ranging from high-temperature superconductivity to Mott insulation[4] — the electronic version of a traffic jam. An ongoing research problem is to determine how these strong interactions and their weakening through laser driving[5] modify high-harmonic spectra[6,7]. Lakhotia and colleagues' paper could be seen as motivation to search for a path towards imaging such strong electron–electron interactions.

Finally, a key direction for future work concerns the dynamic imaging of the interplay between driven electrons and other excitations in strongly driven quantum materials, in particular at even longer laser wavelengths than those used in this study. The first step towards this goal is the reconstruction of interatomic potential-energy landscapes from highly displaced nuclei[8]. It will be intriguing to see how the combination of different time-domain techniques will provide a glimpse into the complex interplay of the many constituents from which fascinating material properties emerge in and out of equilibrium[9].

Michael A. Sentef is at the Max Planck Institute for the Structure and Dynamics of Matter, 22761 Hamburg, Germany and the Institute for Theoretical Physics, University of Bremen, Bremen, Germany.
e-mail: michael.sentef@mpsd.mpg.de

1. Lakhotia, H. *et al. Nature* **583**, 55–59 (2020).
2. Morales, F., Richter, M., Patchkovskii, S. & Smirnova, O. *Proc. Natl Acad. Sci. USA* **108**, 16906–16911 (2011).
3. Ghimire, S. & Reis, D. A. *Nature Phys.* **15**, 10–16 (2019).
4. Keimer, B., Kivelson, S. A., Norman, M. R., Uchida, S. & Zaanen, J. *Nature* **518**, 179–186 (2015).
5. Tancogne-Dejean, N., Sentef, M. A. & Rubio, A. *Phys. Rev. Lett.* **121**, 097402 (2018).
6. Silva, R. E. F., Blinov, I. V., Rubtsov, A. N., Smirnova, O. & Ivanov, M. *Nature Photon.* **12**, 266–270 (2018).
7. Murakami, Y., Eckstein, M. & Werner, P. *Phys. Rev. Lett.* **121**, 057405 (2018).
8. von Hoegen, A., Mankowsky, R., Fechner, M., Först, M. & Cavalleri, A. *Nature* **555**, 79–82 (2018).
9. Gerber, S. *et al. Science* **357**, 71–75 (2017).

Immunology

T cells engineered to target senescence

Verena Wagner & Jesús Gil

Senescence is a hallmark of cellular ageing and contributes to many diseases. A new method enabling immune cells to target senescent cells might offer improved therapeutic options. **See p.127**

Senescence is a form of cellular stress response. In some circumstances it can be harmful, and efforts are under way to develop therapies that target senescent cells. On page 127, Amor *et al.*[1] describe a method that selectively removes senescent cells in mice.

Entry into senescence imposes a stable arrest of the cell cycle, preventing old, damaged or precancerous cells from dividing. Senescent cells secrete a complex cocktail of factors that drive a response called the senescence-associated secretory phenotype (SASP). This recruits T cells and NK cells of the immune system, promoting removal of the senescent cells. Under these conditions, senescence is transient, which benefits the organism[2].

However, when senescent cells linger, they can promote chronic inflammation resulting in age-related diseases such as atherosclerosis, cancer and fibrosis (a type of tissue scarring). The elimination of senescent cells has therefore emerged as a promising therapeutic strategy. It can improve the outcome of many diseases, and increases lifespan in studies in mice[3]. One possible way to target senescent cells is with drugs that kill them selectively, called senolytic drugs. Amor and colleagues take a different approach, inspired by the fact that immune cells are involved in eliminating senescent cells under normal circumstances[4].

The authors adapted a technique that is currently in use for anticancer treatment. In this therapy, T cells are removed from an individual and, before being returned, are manipulated to boost their ability to target cancer cells. Such cells are known as CAR T cells because they are engineered to express what is termed a chimaeric antigen receptor (CAR). The CAR is designed to recognize and bind to a particular fragment of a protein, called an antigen, that is present on the surface of cancer cells. If this interaction occurs, the T cell is activated and kills the tumour cells[5]. Identifying antigens that are expressed exclusively on tumour cells is a key challenge, because the killing of healthy cells by CAR T cells could lead to severe side effects.

To find antigens that are specific to senescent cells, Amor and colleagues analysed the expression of transmembrane proteins found in senescent human and mouse cells. One of the eight most promising candidates identified was the urokinase-type plasminogen activator receptor (uPAR). An examination of previously published data on protein and RNA expression in human tissues revealed that uPAR is either not detected or is present only at low levels in most organs of the human body, including the central nervous system, heart and liver. However, Amor and colleagues found that uPAR is highly expressed in senescent cells both *in vitro* and *in vivo*. Intriguingly, a soluble form of uPAR (suPAR) that lacks a transmembrane region is a component secreted during the SASP response. The presence of suPAR is a hallmark of some chronic disorders, including diabetes[6] and kidney disease[7], in which senescence has a role.

After identifying uPAR as a universal marker of senescent cells, Amor and colleagues engineered CAR T cells to target uPAR (Fig. 1). Given that premalignant cells (those possibly on their way to becoming cancer cells) undergo senescence, and the fact that many anticancer therapies work by causing tumour cells to enter senescence as a way of stopping them dividing, the authors investigated how effective these CAR T cells were in treating cancer. They report that treatment with CAR T cells that target uPAR eliminated senescent premalignant and malignant cells in mouse models of liver and lung cancer. It has already been proposed that anticancer therapies might be improved by following them up with treatments targeting senescent cells[8]. Amor and colleagues' study in mice confirmed that such an approach using their senolytic CAR T cells boosts the effectiveness of anticancer treatment.

Part of the attraction of using senolytic CAR T cells is their potential for treating the many diseases in which senescence is involved. Indeed, Amor and colleagues show that if mice received senolytic CAR T cells, this improved the outcome of liver fibrosis in animal models of non-alcoholic steatohepatitis, a severe form of fatty liver disease.

Navitoclax, a senolytic drug widely used in preclinical studies, can cause toxicity that limits its use. This has led to efforts to identify new senolytic drugs and other ways to target

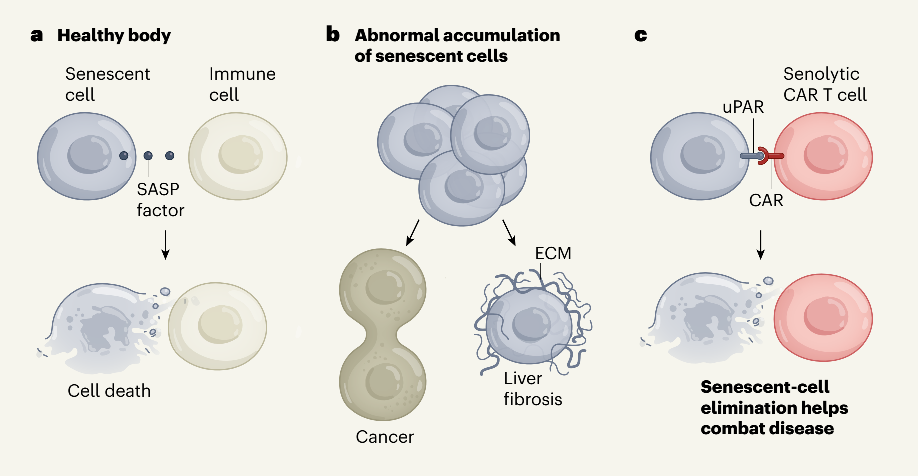


Figure 1 | CAR T cells can be used to remove senescent cells. a, Malfunctioning cells commonly enter a non-dividing state called senescence. These cells exhibit a response called the senescence-associated secretory phenotype (SASP). This is associated with the release of various molecules that attract immune cells, which then kill the senescent cell. **b**, If this process fails and senescent cells linger, they can contribute to diseases such as cancer and liver fibrosis – tissue scarring associated with deposits of extracellular matrix (ECM) material. **c**, Amor *et al.*[1] describe a method that selectively removes senescent cells. The authors identified a protein (uPAR) that is expressed on the surface of senescent cells, and engineered immune cells called T cells to express a receptor that recognizes uPAR. This type of receptor is called a chimaeric antigen receptor (CAR). The recognition process drives the T cells to kill the senescent cells. Amor and colleagues report that such senolytic CAR T cells help to tackle disease in mouse models of cancer and liver fibrosis.

senescent cells[9]. Amor *et al*. suggest that use of senolytic CAR T cells could eliminate some of the side effects and the limited effectiveness associated with senolytic drugs. However, these cells are not necessarily problem-free. A common complication of the therapeutic use of CAR T cells is a condition called cytokine-release syndrome (also known as a cytokine storm), in which an intense T-cell response causes fever and affects blood pressure and breathing[10]. Although the authors observed that high doses of senolytic CAR T cells did cause cytokine-release syndrome, reducing the dosage avoided the problem while retaining the therapeutic potential of the treatment.

The use of CAR T cells for anticancer therapy has other limitations. Long-lasting activity of these cells is required to control tumour growth as cancer cells divide over time. This issue might not be of concern when targeting senescent cells, because they do not proliferate. However, many solid tumours (those that do not arise from blood cells) are associated with an immunosuppressive tissue microenvironment, which can cause CAR T cells to enter a dysfunctional state called exhaustion. Senescent cells can foster an immunosuppressive microenvironment during tumour formation[11]. Although the authors did not observe senescence-mediated immunosuppression in their study, it might be a shortcoming of this approach. A greater understanding is needed of how senescent cells can interfere with immune-system function.

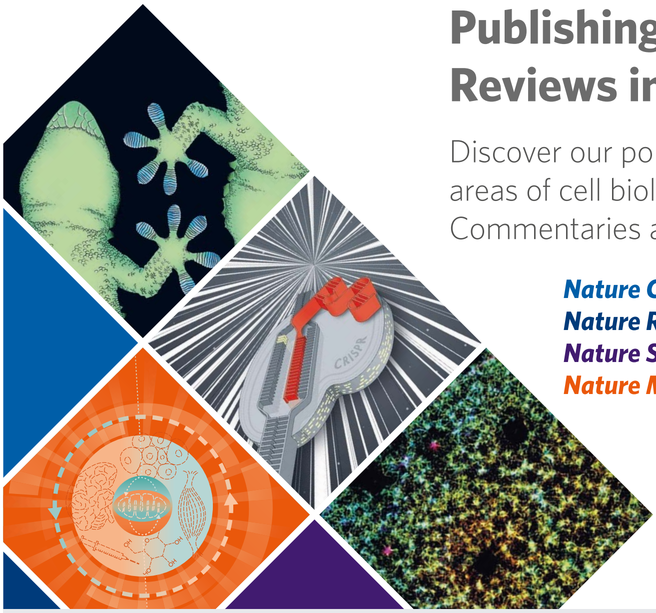
Could senolytic CAR T cells be used to treat patients? The use of such cells in the clinic is expensive, so the criteria for considering such an approach should be chosen carefully. It will also be important to determine whether CAR T cells that target human uPAR are as safe and effective as are the CAR T cells targeting mouse uPAR that Amor and colleagues used. Alternatively, perhaps this method could be improved by using senolytic CAR T cells that target other proteins found on the surface of senescent cells, such as DPP4 and oxidized vimentin[12,13]. The immense advances that are being made in mapping gene expression in humans at the resolution of single cells might reveal further targets for use in the design of senolytic CAR T cells. Merging two promising therapeutic strategies by using CAR T cells to target senescent cells might be a powerful combination for tackling certain diseases.

> **“When senescent cells linger, they can promote chronic inflammation resulting in age-related diseases.”**

Verena Wagner and **Jesús Gil** are at the MRC London Institute of Medical Sciences, London W12 0NN, UK, and at the Institute of Clinical Sciences, Faculty of Medicine, Imperial College London.
e-mail: jesus.gil@imperial.ac.uk

1. Amor, C. *et al. Nature* 583, 127–132 (2020).
2. Gorgoulis, V. *et al. Cell* **179**, 813–827 (2019).
3. Baker, D. J. *et al. Nature* **530**, 184–189 (2016).
4. Kang, T.-W. *et al. Nature* **479**, 547–551 (2011).
5. Lim, W. A. & June, C. H. *Cell* **168**, 724–740 (2017).
6. Curovic, V. R. *et al. Diabetes Care* **42**, 1112–1119 (2019).
7. Quaglia, M., Musetti, C. & Cantaluppi, V. *N. Engl. J. Med.* **374**, 890–891 (2016).
8. Wang, L. & Bernards, R. *Front. Med.* **12**, 490–495 (2018).
9. Ovadya, Y. & Krizhanovsky, V. *J. Clin. Invest.* **128**, 1247–1254 (2018).
10. Brudno, J. N. & Kochenderfer, J. N. *Blood* **127**, 3321–3330 (2016).
11. Eggert, T. *et al. Cancer Cell* **30**, 533–547 (2016).
12. Kim, K. M. *et al. Genes Dev.* **31**, 1529–1534 (2017).
13. Frescas, D. *et al. Proc. Natl Acad. Sci. USA* **114**, E1668–E1677 (2017).

J.G. declares competing financial interests: see go.nature.com/3crv3as for details.



A remnant planetary core in the hot-Neptune desert

https://doi.org/10.1038/s41586-020-2421-7

Received: 29 October 2019

Accepted: 20 March 2020

A list of authors and their affililations appears at the end of the paper.

The interiors of giant planets remain poorly understood. Even for the planets in the Solar System, difficulties in observation lead to large uncertainties in the properties of planetary cores. Exoplanets that have undergone rare evolutionary processes provide a route to understanding planetary interiors. Planets found in and near the typically barren hot-Neptune 'desert'[1,2] (a region in mass–radius space that contains few planets) have proved to be particularly valuable in this regard. These planets include HD149026b[3], which is thought to have an unusually massive core, and recent discoveries such as LTT9779b[4] and NGTS-4b[5], on which photoevaporation has removed a substantial part of their outer atmospheres. Here we report observations of the planet TOI-849b, which has a radius smaller than Neptune's but an anomalously large mass of $39.1^{+2.7}_{-2.6}$ Earth masses and a density of $5.2^{+0.7}_{-0.8}$ grams per cubic centimetre, similar to Earth's. Interior-structure models suggest that any gaseous envelope of pure hydrogen and helium consists of no more than $3.9^{+0.8}_{-0.9}$ per cent of the total planetary mass. The planet could have been a gas giant before undergoing extreme mass loss via thermal self-disruption or giant planet collisions, or it could have avoided substantial gas accretion, perhaps through gap opening or late formation[6]. Although photoevaporation rates cannot account for the mass loss required to reduce a Jupiter-like gas giant, they can remove a small (a few Earth masses) hydrogen and helium envelope on timescales of several billion years, implying that any remaining atmosphere on TOI-849b is likely to be enriched by water or other volatiles from the planetary interior. We conclude that TOI-849b is the remnant core of a giant planet.

The TESS mission[7] observed the star TOI-849/TIC33595516 (V magnitude 12) for 27 days during September and October 2018, leading to the detection of a candidate transiting planet. TOI-849 was observed at a cadence of 30 min in the full-frame images using the MIT Quick Look pipeline (see Methods). No signs of additional planets or stellar activity were seen in the photometry. Follow-up observations with the HARPS spectrograph detected a large radial velocity signal, confirming the planet TOI-849b. Four additional transits were observed using the ground-based telescopes Next Generation Transit Survey[8] and Las Cumbres Observatory Global Telescope[9], improving the radius determination and ephemeris of the planet. A search in Gaia Data Release 2 reveals no other sources closer than 39″, with the closest source being 7.8 magnitudes fainter than TOI-849 in the G band[10]. Additional high-resolution imaging from SOAR, NACO/VLT, AstraLux and Zorro/Gemini South revealed no unresolved companion stars. We perform a joint fit to the data using the PASTIS software[11] to extract planetary and stellar parameters, using the combined HARPS spectra to derive priors on the stellar parameters and calculate chemical abundances for the host star (see Methods). The best fit and data are shown in Fig. 1.

TOI-849b has a mass of $39.1^{+2.7}_{-2.6}M_{\oplus}$ ($M_{\oplus}$, mass of Earth), nearly half the mass of Saturn (all uncertainties are 1σ unless otherwise stated). The planet's radius is $3.44^{+0.16}_{-0.12}R_{\oplus}$ ($R_{\oplus}$, Earth radius) and its mean density is $5.2^{+0.7}_{-0.8}$ g cm^{-3}, making it the densest Neptune-sized planet discovered so far (Fig. 2). It has an orbital period of 0.7655241 ± 0.0000027 d, making it an 'ultrashort-period' planet. The upper limit on its eccentricity is 0.08 at 95% confidence. Its radius, mass and period place TOI-849b in the middle of the hot-Neptune desert, a region of parameter space typically devoid of planets due to photoevaporation and tidal disruption[1,2] (Fig. 3). The host star TOI-849 is a late G dwarf with mass $(0.929 \pm 0.023)M_{\odot}$, radius $0.919^{+0.029}_{-0.023}R_{\odot}$ ($M_{\odot}$ and $R_{\odot}$ are the mass and radius of the Sun, respectively) and age $6.7^{+2.9}_{-2.4}$ Gyr. The close proximity of the planet and the star lead to an equilibrium temperature of 1,800 K for the planet, assuming an albedo of 0.3. The full set of derived parameters for the planet and star are given in Extended Data Tables 1, 2 and general stellar parameters are provided in Extended Data Table 3.

The most widely used interior-structure models for terrestrial planets are not valid for planets as massive as TOI-849b, because the properties of matter at such high central pressures remain highly uncertain. Furthermore, some compositional mixing is expected at these high pressures and temperatures[12], in contradiction of the usual assumption of distinct layers[13]. We build an internal-structure model accounting for some of these issues (see Methods), but restrict our analysis to the limiting cases of a maximum and minimum possible hydrogen and helium (H/He) envelope mass under the layered-structure assumption. We calculate the maximum envelope mass by minimizing the contribution of the core, mantle and water, assuming that the planet has the same [Fe/Si] ratio as that observed for the photosphere of the host star. Under this model, the maximum envelope mass fraction is $3.9^{+0.8}_{-0.9}$%.

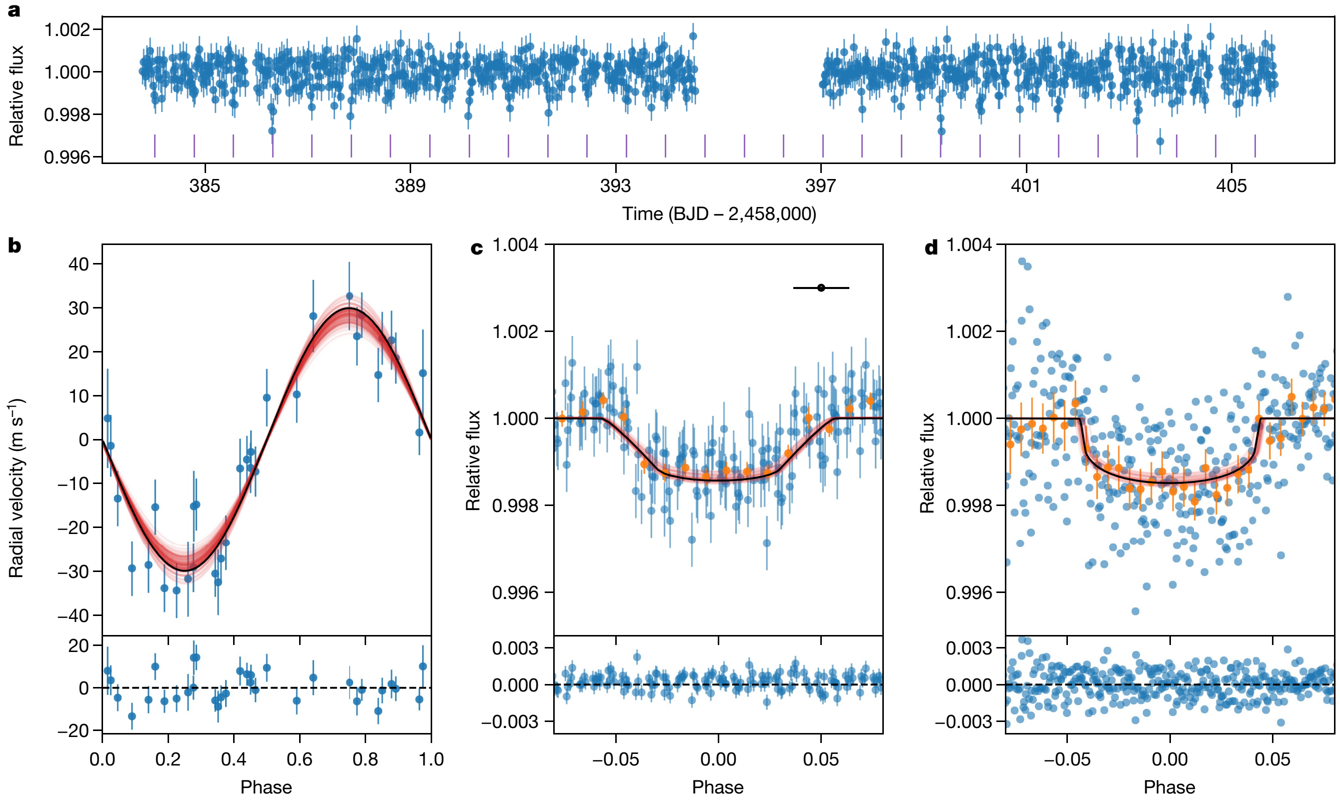


Fig. 1 | Best-fitting model to the TESS, HARPS and NGTS data. a, TESS light curve with transit times marked as vertical lines. BJD, barycentric Julian date. **b**, Phase-folded HARPS data (blue symbols) and best-fitting model (black line), with residuals shown in the bottom panel. Several models randomly drawn from the Markov chain Monte Carlo output are shown in red. **c**, Phase-folded TESS 30-min cadence data (blue symbols) and binned to 0.01 in phase (orange symbols, nine individual points per bin), with models as in **b** and residuals shown in the bottom panel. The horizontal error bar shows the TESS cadence. **d**, Phase-folded NGTS data binned to 1 min (blue symbols, 46 individual points per bin) and to 0.01 in phase (orange symbols, 777 individual points per bin). We plot the binned NGTS data to aid visualization but we fit the models to the full dataset. Models are as in **b**, with residuals in the bottom panel. The cadence is negligible at this scale. Data from Las Cumbres Observatory Global Telescope were also used and are shown in Extended Data Fig. 1. Vertical error bars of individual points show one standard deviation. In the case of binned measurements, points and error bars show the weighted mean and its standard error, respectively.

The large core mass and low envelope mass fraction of TOI-849b challenge the traditional view of planet formation via core accretion, in which planets with masses above a critical mass of about $10M_{\oplus}$–$20M_{\oplus}$ are expected to undergo runaway gas accretion within the protoplanetary disk[14–16]. Why, then, does TOI-849b lack a massive gaseous envelope? Apparently the core somehow avoided runaway accretion, or else the planet was once a gas giant that somehow lost most of its envelope. If runaway accretion proceeded to produce a giant planet, removal of most of the original mass would be required to reach the present-day state. HD149026b[3] is a giant planet with mass $(121 \pm 19)M_{\oplus}$ (ref. [17]), which is thought to have a solid core with a mass of about $50M_{\oplus}$ (refs. [18,19]), similar to TOI-849b. Starting from a planet such as HD149026b, a mass loss of 60–70% would be required to produce the present-day TOI-849b. Considering the proximity of TOI-849b to its host star, one would expect some mass loss to photoevaporation. The predicted lifetime mass-loss rate for a Jupiter-like planet is only a few per cent, well below the required range (see Methods). For a planet such as HD149026b, the situation is less clear, and the lifetime mass removed depends critically on the assumptions made. We proceed to explore several formation pathways for TOI-849b.

Tidal disruption could cause a mass loss of 1–2 orders of magnitude. The close proximity of a number of hot Jupiters to their tidal disruption radii[20] and the fact that hot Jupiters are preferentially found around younger stars[21,22] suggest that tidal disruption of hot Jupiters might be frequent. Although it appears that they do not typically leave behind a remnant core, or such cores are short-lived[23], as a rare higher-mass object, TOI-849b may be an unusual case. At the location of TOI-849b, tidal disruption would be expected for a Jupiter-mass planet with radius greater than 1.5 Jupiter radii. An alternative related pathway to substantial envelope loss is disruption via tidal thermalization events, which can lead to mass loss of 1–2 orders of magnitude. If TOI-849b reached its close orbit via high-eccentricity scattering by another planet in the system, energy buildup in the planet's internal f-modes during tidal circularization could approach large fractions of the planet's internal binding energy and potentially lead to thermalization events, which may remove envelope layers (see Methods). However, in either case it is unclear whether a giant planet could harbour a large enough core to leave behind a $40M_{\oplus}$ remnant, because the gaseous envelope on top of a core of a few Earth masses causes planetesimals to be eroded in the envelope. The remaining solids must subsequently rain out to produce such a large core[12,24,25].

Giant planet collisions provide another, intermediate way to produce planets similar to TOI-849b. The Bern planetary population synthesis models[26] predict the existence of a small population of planets with similar masses and semi-major axes to TOI-849b (see Methods). In those models, such planets are produced via giant planet collisions at the end of the migration phase, resulting in the ejection of the planetary envelope, leaving no time for the remnant core to accrete further gas. In these scenarios, the cores reach an envelope mass fraction of a few tens of per cent before being reduced to Neptune size and ejecting the envelope through an impact. Such a scenario leaves a dense planetary core close to the host star.

The alternative hypothesis is that TOI-849b avoids runaway accretion, possibly through opening a gap in the protoplanetary disk that is largely devoid of gas, before the planet accretes much envelope mass.

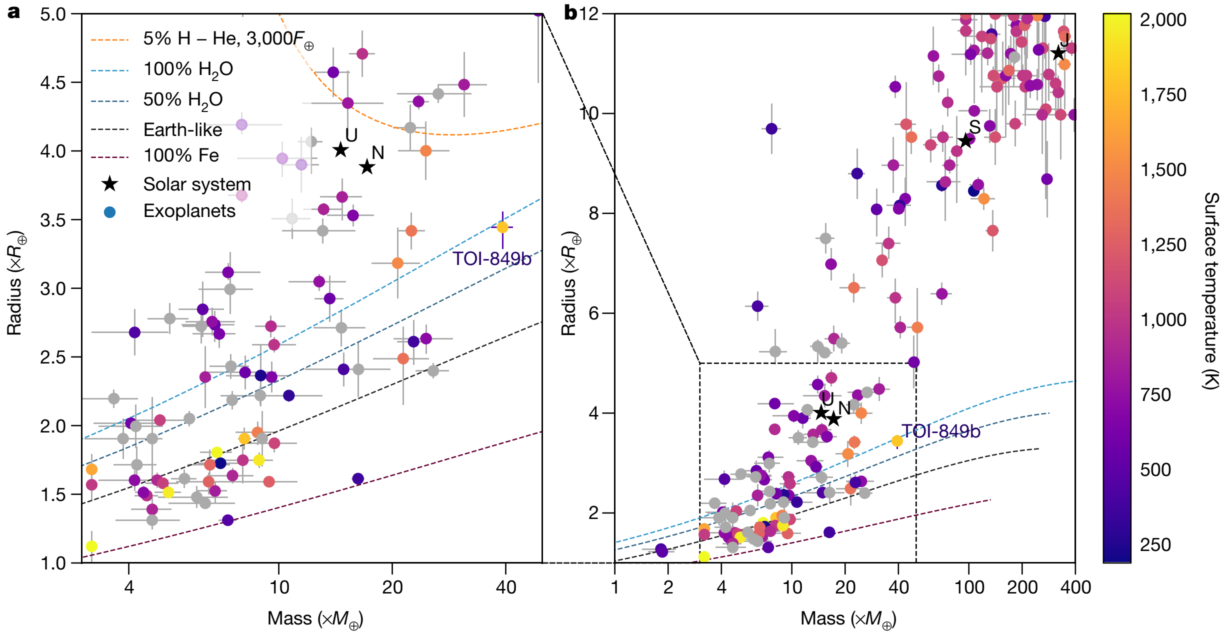


Fig. 2 | Mass–radius diagram of known exoplanets from the NASA exoplanet archive. a, **b**, The archive (https://exoplanetarchive.ipac.caltech.edu/) was accessed on 20 January 2020. Planets are coloured according to calculated equilibrium temperature and are grey otherwise. Planets with mass determinations better than 4σ are shown. Planets without a reported mass determination were excluded[30]. Composition tracks[31] are shown as dashed lines, with an additional 5% H–He track at an irradiation level similar to that of TOI-849b. U, N, S and J denote the Solar System planets Uranus, Neptune, Saturn and Jupiter, respectively. $F_\oplus$ represents the average solar irradiation received by Earth. **a**, Zoom of **b**. All error bars show one standard deviation.

Because the threshold mass required for a planet to open up a gap in a protoplanetary disk is sensitive to the disk scale height, which is small close to the star, planets on close-in orbits can more easily open a deep gap. A 40$M_\oplus$ planet such as TOI-849b on an orbit of 0.1 AU would reduce the disk surface density at its location by a factor about 10 (refs. [27,28]). Recently, it has been argued that a reduction in gas accretion due to gap opening is required to resolve the fact that runaway gas accretion models tend to produce too many Jupiter-mass planets and not enough sub-Saturn-mass planets[6]. Indeed, by reducing the accretion rate onto gap-opening planets, it is possible to produce 40$M_\oplus$ planets at 0.1 AU with gas mass fractions below 10% if the planets form late enough[6]. In contrast to the tidal disruption pathway, reduced gas accretion should leave TOI-849b aligned with the stellar spin axis. Detecting or ruling out such alignment using measurements of the Rossiter–McLaughlin effect[29], as well as taking measurements of the atmospheric composition, may aid in distinguishing between the various formation scenarios.

In all cases, remaining hydrogen and helium envelope masses of a few per cent could be removed over several billion years by photoevaporation, given the planet's close orbit. We estimate the current mass-loss rate to be 0.95$M_\oplus$ Gyr^{-1} (see Methods), which implies that an envelope mass of ~4% could be removed in a few billion years. Therefore, the question changes: where does TOI-849b's minor envelope come from? Given the high equilibrium temperature, we would expect that some ices would be evaporated to provide a secondary enriched atmosphere containing water and other volatiles. In these circumstances, TOI-849b provides a unique target in which the composition of a primordial planetary core could be studied by observing its atmospheric constituents with, for example, the Hubble Telescope or the upcoming James Webb Space Telescope.

The proximity of TOI-849b to its host star, which promotes gap opening and increases the role of photoevaporation, could explain why similar objects have not yet been found. Ultimately, however TOI-849b formed, the planet's large mass and low gas mass fraction will provide a stringent test of planet formation theory. TOI-849b gives us a glimpse at a core similar to those that exist at the centres of giant planets, exposed through an unlikely combination of inhibited accretion or mass loss.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2421-7.

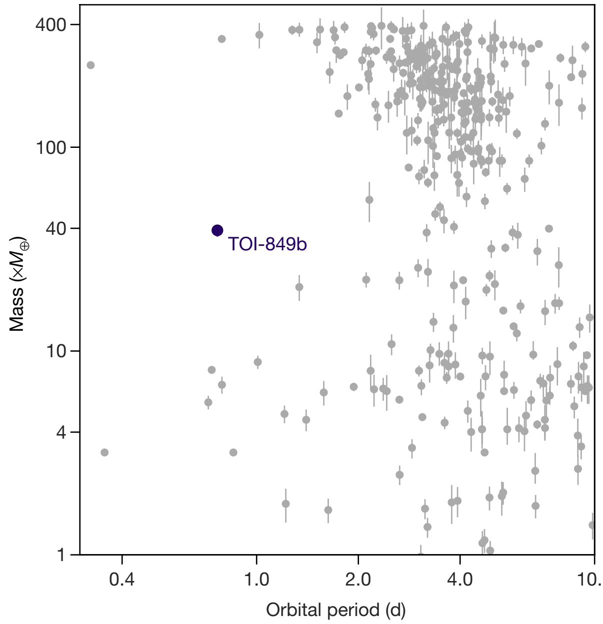


Fig. 3 | TOI-849b in the context of the hot-Neptune desert. Known exoplanets are plotted in grey and were sourced from the NASA exoplanet archive (https://exoplanetarchive.ipac.caltech.edu/) on 20th January 2020. Only planets with mass determinations better than 4σ are plotted. All error bars show one standard deviation.

1. Szabó, G. M. & Kiss, L. L. A short-period censor of sub-Jupiter mass exoplanets with low density. *Astrophys. J. Lett.* **727**, L44 (2011).
2. Owen, J. E. & Lai, D. Photoevaporation and high-eccentricity migration created the sub-Jovian desert. *Mon. Not. R. Astron. Soc.* **479**, 5012–5021 (2018).
3. Sato, B. et al. The N2K Consortium. II. A transiting hot Saturn around HD 149026 with a large dense core. *Astrophys. J.* **633**, 465–473 (2005).
4. Jenkins, J. TESS Discovery of the first ultra hot Neptune, LTT9779b. In *AAS/Division for Extreme Solar Systems Abstracts* Vol. 51, 103.07 (American Astronomical Society, 2019).
5. West, R. G. et al. NGTS-4b: a sub-Neptune transiting in the desert. *Mon. Not. R. Astron. Soc.* **486**, 5094–5103 (2019).
6. Lee, E. J. The boundary between gas-rich and gas-poor planets. *Astrophys. J.* **878**, 36 (2019).
7. Ricker, G. R. et al. Transiting exoplanet survey satellite (TESS). *J. Astron. Telesc. Instrum. Syst.* **1**, 014003 (2014).
8. Wheatley, P. J. et al. The next generation transit survey (NGTS). *Mon. Not. R. Astron. Soc.* **475**, 4476–4493 (2018).
9. Brown, T. M. et al. Las Cumbres Observatory global telescope network. *Publ. Astron. Soc. Pacific* **125**, 1031 (2013).
10. Gaia Collaboration. Gaia Data Release 2. Summary of the contents and survey properties. *Astron. Astrophys.* **616**, A1 (2018).
11. Santerne, A. et al. PASTIS: Bayesian extrasolar planet validation - II. Constraining exoplanet blend scenarios using spectroscopic diagnoses. *Mon. Not. R. Astron. Soc.* **451**, 2337–2351 (2015).
12. Bodenheimer, P., Stevenson, D. J., Lissauer, J. J. & D'Angelo, G. New formation models for the Kepler-36 system. *Astrophys. J.* **868**, 138 (2018).
13. Dorn, C. et al. A generalized Bayesian inference method for constraining the interiors of super Earths and sub-Neptunes. *Astron. Astrophys.* **597**, A37 (2017).
14. Mizuno, H., Nakazawa, K. & Hayashi, C. Instability of a gaseous envelope surrounding a planetary core and formation of giant planets. *Prog. Theor. Phys.* **60**, 699–710 (1978).
15. Rafikov, R. R. Atmospheres of protoplanetary cores: critical mass for nucleated instability. *Astrophys. J.* **648**, 666–682 (2006).
16. Piso, A.-M. A., Youdin, A. N. & Murray-Clay, R. A. Minimum core masses for giant planet formation with realistic equations of state and opacities. *Astrophys. J.* **800**, 82 (2015).
17. Stassun, K. G., Collins, K. A. & Gaudi, B. S. Accurate empirical radii and masses of planets and their host stars with gaia parallaxes. *Astron. J.* **153**, 136 (2017).
18. Fortney, J. J., Saumon, D., Marley, M. S., Lodders, K. & Freedman, R. S. Atmosphere, interior, and evolution of the metal-rich transiting planet HD 149026b. *Astrophys. J.* **642**, 495–504 (2006).
19. Ikoma, M., Guillot, T., Genda, H., Tanigawa, T. & Ida, S. On the origin of HD 149026b. *Astrophys. J.* **650**, 1150–1159 (2006).
20. Delrez, L. et al. WASP-121 b: a hot Jupiter close to tidal disruption transiting an active F star. *Mon. Not. R. Astron. Soc.* **458**, 4025–4043 (2016).
21. Collier Cameron, A. & Jardine, M. Hierarchical Bayesian calibration of tidal orbit decay rates among hot Jupiters. *Mon. Not. R. Astron. Soc.* **476**, 2542–2555 (2018).
22. Hamer, J. H. & Schlaufman, K. C. Hot Jupiters are destroyed by tides while their host stars are on the main sequence. *Astron. J.* **158**, 190 (2019).
23. Winn, J. N. et al. Absence of a metallicity effect for ultra-short-period planets. *Astron. J.* **154**, 60 (2017).
24. Iaroslavitz, E. & Podolak, M. Atmospheric mass deposition by captured planetesimals. *Icarus* **187**, 600–610 (2007).
25. Brouwers, M. G., Vazan, A. & Ormel, C. W. How cores grow by pebble accretion. I. Direct core growth. *Astron. Astrophys.* **611**, A65 (2018).
26. Mordasini, C. in *Handbook of Exoplanets* (ed. Pudritz, R.) 2425–2474 (Springer, 2018).
27. Crida, A., Morbidelli, A. & Masset, F. On the width and shape of gaps in protoplanetary disks. *Icarus* **181**, 587–604 (2006).
28. Duffell, P. C. & MacFadyen, A. I. Gap opening by extremely low-mass planets in a viscous disk. *Astrophys. J.* **769**, 41 (2013).
29. Triaud, A. H. M. J. in *Handbook of Exoplanets* (eds Deeg, H. & Belmonte, J.) 1375–1401 (Springer, 2018).
30. Xie, J.-W. Transit timing variation of near-resonance planetary pairs. II. Confirmation of 30 planets in 15 multiple-planet systems. *Astrophys. J. Suppl. Ser.* **210**, 25 (2014).
31. Zeng, L. & Sasselov, D. A detailed model grid for solid planets from 0.1 through 100 Earth masses. *Publ. Astron. Soc. Pacif.* **125**, 227 (2013).

David J. Armstrong[1,2]✉, Théo A. Lopez[3], Vardan Adibekyan[4], Richard A. Booth[5], Edward M. Bryant[1,2], Karen A. Collins[6], Magali Deleuil[3], Alexandre Emsenhuber[7,8], Chelsea X. Huang[9], George W. King[1,2], Jorge Lillo-Box[10], Jack J. Lissauer[11], Elisabeth Matthews[9], Olivier Mousis[3], Louise D. Nielsen[12], Hugh Osborn[3], Jon Otegi[12,13], Nuno C. Santos[4,14], Sérgio G. Sousa[4], Keivan G. Stassun[15,16], Dimitri Veras[1,2], Carl Ziegler[17], Jack S. Acton[18], Jose M. Almenara[19], David R. Anderson[1,2], David Barrado[10], Susana C. C. Barros[4], Daniel Bayliss[1,2], Claudia Belardi[18], Francois Bouchy[12], César Briceño[20], Matteo Brogi[1,2,21], David J. A. Brown[1,2], Matthew R. Burleigh[18], Sarah L. Casewell[18], Alexander Chaushev[22], David R. Ciardi[23], Kevin I. Collins[24], Knicole D. Colón[25], Benjamin F. Cooke[1,2], Ian J. M. Crossfield[9], Rodrigo F. Díaz[26,27,28], Elisa Delgado Mena[4], Olivier D. S. Demangeon[4], Caroline Dorn[13], Xavier Dumusque[12], Philipp Eigmüller[29], Michael Fausnaugh[9], Pedro Figueira[4,30], Tianjun Gan[31], Siddharth Gandhi[2], Samuel Gill[1,2], Erica J. Gonzales[32], Michael R. Goad[18], Maximilian N. Günther[9], Ravit Helled[13], Saeed Hojjatpanah[4,14], Steve B. Howell[11], James Jackman[1,2], James S. Jenkins[33,34], Jon M. Jenkins[11], Eric L. N. Jensen[35], Grant M. Kennedy[1,2], David W. Latham[36], Nicholas Law[37], Monika Lendl[12,38], Michael Lozovsky[13], Andrew W. Mann[37], Maximiliano Moyano[39], James McCormac[1,2], Farzana Meru[1,2], Christoph Mordasini[8], Ares Osborn[1,2], Don Pollacco[1,2], Didier Queloz[40], Liam Raynard[18], George R. Ricker[9], Pamela Rowden[41], Alexandre Santerne[3], Joshua E. Schlieder[25], Sara Seager[9,42,43], Lizhou Sha[9], Thiam-Guan Tan[44], Rosanna H. Tilbrook[18], Eric Ting[11], Stéphane Udry[12], Roland Vanderspek[9], Christopher A. Watson[45], Richard G. West[1,2], Paul A. Wilson[1,2], Joshua N. Winn[46], Peter Wheatley[1,2], Jesus Noel Villasenor[9], Jose I. Vines[33] & Zhuchang Zhan[42]

---

[1]Centre for Exoplanets and Habitability, University of Warwick, Coventry, UK. [2]Department of Physics, University of Warwick, Coventry, UK. [3]Aix-Marseille Université, CNRS, CNES, LAM, Marseilles, France. [4]Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, CAUP, Porto, Portugal. [5]Institute of Astronomy, University of Cambridge, Cambridge, UK. [6]Center for Astrophysics, Harvard and Smithsonian, Cambridge, MA, USA. [7]Lunar and Planetary Laboratory, University of Arizona, Tucson, AZ, USA. [8]Physikalisches Institut, University of Bern, Bern, Switzerland. [9]Department of Physics and Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA, USA. [10]Departamento de Astrofísica, Centro de Astrobiología (CSIC-INTA), Madrid, Spain. [11]NASA Ames Research Center, Moffett Field, CA, USA. [12]Observatoire Astronomique de l'Université de Genéve, Versoix, Switzerland. [13]Institute for Computational Science, University of Zurich, Zurich, Switzerland. [14]Departamento de Física e Astronomia, Faculdade de Ciências, Universidade do Porto, Porto, Portugal. [15]Department of Physics and Astronomy, Vanderbilt University, Nashville, TN, USA. [16]Department of Physics, Fisk University, Nashville, TN, USA. [17]Dunlap Institute for Astronomy and Astrophysics, University of Toronto, Toronto, Ontario, Canada. [18]School of Physics and Astronomy, University of Leicester, Leicester, UK. [19]Université Grenoble Alpes, CNRS, IPAG, Grenoble, France. [20]Cerro Tololo Inter-American Observatory, La Serena, Chile. [21]INAF Osservatorio Astrofisico di Torino, Pino Torinese, Italy. [22]Center for Astronomy and Astrophysics, TU Berlin, Berlin, Germany. [23]Caltech/IPAC-NASA Exoplanet Science Institute, Pasadena, CA, USA. [24]George Mason University, Fairfax, VA, USA. [25]NASA Goddard Space Flight Center, Exoplanets and Stellar Astrophysics Laboratory (Code 667), Greenbelt, MD, USA. [26]Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina. [27]Instituto de Astronomía y Física del Espacio (IAFE), CONICET – Universidad de Buenos Aires, Buenos Aires, Argentina. [28]International Center for Advanced Studies (ICAS) and ICIFI (CONICET), ECyT-UNSAM, Campus Miguelente, Buenos Aires, Argentina. [29]Institute of Planetary Research, German Aerospace Center, Berlin, Germany. [30]European Southern Observatory, Santiago, Chile. [31]Department of Astronomy and Tsinghua Centre for Astrophysics, Tsinghua University, Beijing, China. [32]Department of Astronomy and Astrophysics, University of California at Santa Cruz, Santa Cruz, CA, USA. [33]Departamento de Astronomía, Universidad de Chile, Santiago, Chile. [34]Centro de Astrofísica y Tecnologías Afines (CATA), Santiago, Chile. [35]Department of Physics and Astronomy, Swarthmore College, Swarthmore, PA, USA. [36]Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA. [37]Department of Physics and Astronomy, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [38]Space Research Institute, Austrian Academy of Sciences, Graz, Austria. [39]Instituto de Astronomía, Universidad Católica del Norte, Antofagasta, Chile. [40]Cavendish Laboratory, University of Cambridge, Cambridge, UK. [41]School of Physical Sciences, The Open University, Milton Keynes, UK. [42]Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA. [43]Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA, USA. [44]Perth Exoplanet Survey Telescope, Perth, Western Australia, Australia. [45]Astrophysics Research Centre, Queen's University Belfast, Belfast, UK. [46]Department of Astrophysical Sciences, Princeton University, Princeton, NJ, USA. ✉e-mail: d.j.armstrong@warwick.ac.uk

Methods

Observations and analysis

TESS. TOI-849 was observed in TESS (Transiting Exoplanet Survey Satellite) sector 3 (20 September 2018 to 18 October 2018), Camera 2 and CCD 3, with 30-min cadence on the full-frame images (FFIs). The calibrated FFIs, available at the Michulski Archive for Space Telescopes (MAST; https://archive.stsci.edu/missions-and-data/transiting-exoplanet-survey-satellite-tess), were produced by the TESS Science Processing Operations Center (SPOC)[32]. The candidate was detected by the MIT Quick Look pipeline[33] with a signal-to-noise ratio of 18. It exhibited consistent transit depth in the multi-aperture analysis and appeared to be on target in the difference image analysis. It passed all the vetting criteria set by the TESS Science Office and was released as a TESS Object of Interest.

The aperture showing minimal scatter was found to be circular with a radius of 2.5 pixels, and the background was determined on an annulus with a width of 3 pixels and an inner radius of 4 pixels. We rejected outliers due to spacecraft momentum dump with the quaternion time series provided by the spacecraft data. Further long-timescale trends were removed using a B-spline-based algorithm[34]. No evidence of photometric activity was observed. The light curve was further detrended to remove residual long-term trends using a modified Savitzky–Golay filter[35], whereby a sliding window was used to fit a three-dimensional polynomial function to the data while ignoring outliers. Both flattening operations were carried out ignoring in-transit data points. Data before 2,458,383.78 BJD and after 2,458,405.77 BJD were masked because during that time the TESS operations team carried out several experiments on the attitude control system, causing the jitter profile to differ from normal. Data points between 2,458,394.54 BJD and 2,458,397.0 BJD were masked because of scattered light. The resulting light curve is shown in Fig. 1.

NGTS. Two full transits of TOI-849 were observed on the nights of 8 August 2019 and 11 August 2019 UT (universal time) using the Next Generation Transit Survey (NGTS)[8] at the Paranal Observatory of the European Southern Observatory (ESO) in Chile, which are plotted in Fig. 1. The NGTS facility consists of 12 fully robotic 20-cm telescopes coupled to Andor iKon-L 936 cameras, each with an instantaneous field of view of 8 square degrees and a pixel scale of 5″ per pixel. On both nights, ten NGTS telescopes were used to simultaneously observe the transit. Becauise the photometric noise was found to be uncorrelated between the individual NGTS telescopes, we can combine the light curves to achieve ultrahigh precision photometry for TOI-849. A total of 29,654 images were obtained with an exposure time of 10 s using the custom NGTS filter (520–890 nm). The observations were all obtained at an airmass of $z < 2$ and with photometric observing conditions. The telescope guiding was performed using the DONUTS auto-guiding algorithm[36], which provides sub-pixel-level stability of the target position on the charge-coupled device (CCD). We did not require the use of flat fields during the image reduction, because of the high precision of the auto-guiding. This reduction was performed using a custom aperture photometry pipeline, in which the 100 best comparison stars were selected and ranked on the basis of their proximity to the target star in the parameters of on-sky separation, apparent magnitude and colour. This large number of optimized comparison stars ccould be obtained because of the wide field of view of the NGTS telescopes, and again improved the precision of the NGTS light curves by reducing the presence of correlated noise.

HARPS. We obtained radial velocity (RV) measurements of TOI-849 with the High Accuracy Radial velocity Planet Searcher (HARPS) spectrograph (resolving power $R = 115{,}000$) mounted on the 3.6-m telescope at ESO's La Silla Observatory[37]. Thirty three observations were taken between 28 July 2019 and 28 December 2019 in HAM mode, as part of the NCORES large programme (ID 1102.C-0249). An exposure time of at least 1,200 s was used, giving a signal-to-noise ratio of ~20 per pixel. Typically, the star was observed 2–3 times per night. The data were reduced with the offline data reduction software HARPS pipeline. RV measurements were performed using a weighted cross-correlation function (CCF) method with a G2V template[38,39]. The line bisector (BIS) and the full-width at half-maximum (FWHM) were measured using published methods[40]. No correlation was seen between the RVs and the calculated BIS, FWHM, or CCF contrast ($R < 0.09$ in all cases). The RV measurements are listed in Extended Data Table 4, and the RV data, photometry and best fit are shown in Fig. 1. A jitter of 4.2 m s^{-1} was seen, consistent with the low photometric activity level. The BIS and FWHM are shown in Extended Data Fig. 2. We investigated the CCFs for contributions from unresolved stellar companions by removing Gaussian fits to the individual CCF profiles and studying the residuals (Extended Data Fig. 3). No evidence of additional companions was seen. Finally, we studied the RV residuals and found no evidence of further periodicity, as shown in Extended Data Fig. 3.

LCOGT and PEST. Two full transits of TOI-849 were observed on the nights of 30 July 2019 and 9 August 2019 UT in the i′ band using exposure times of 30 s and 40 s, respectively. Data were taken for an additional night on 14 July 2019 UT, which unfortunately missed the transit relative to the revised ephemeris from our joint fit. The data with the transits are plotted in Extended Data Fig. 1. Both observations used the CTIO node of the Las Cumbres Observatory Global Telescope (LCOGT) 1-m network[9]. We used the TESS Transit Finder, which is a customized version of the Tapir software package[41], to schedule our transit observations. The telescopes are equipped with 4,096 × 4,096 LCO SINISTRO cameras having an image scale of 0.389″ per pixel, resulting in a 26′ × 26′ field of view. The images were calibrated using the standard LCOGT BANZAI pipeline, and the photometric data were extracted using the AstroImageJ software package[42]. The first full transit on 30 July was observed with the telescope in focus and achieved a point spread function FWHM of ~1.6″. Circular apertures with radius 3.1″ were used to extract differential photometry for the target star and for all stars within 2.5′ that were brighter than TESS band magnitude 19. All of the neighbouring stars were excluded as possible sources of the TESS detection, and the event was detected on target. A circular aperture with radius 8″ was used for the other LCOGT observation, which was slightly defocused to an FWHM of ~4″. The nearest star in the Gaia Data Release 2 catalogue is 39″ to the north of TOI-849, so the target star photometric apertures were uncontaminated by known nearby stars.

A full transit was observed on 20 August 2019 UT in the R_c band from the Perth Exoplanet Survey Telescope (PEST) near Perth, Australia. The 0.3-m telescope is equipped with a 1,530 × 1,020 SBIG ST-8XME camera with an image scale of 1.2″ per pixel, resulting in a 31′ × 21′ field of view. Systematics at the level of the shallow transit depth precluded inclusion of these data in the joint fit.

NACO/VLT. TOI-849 was imaged with the NAOS/CONICA instrument onboard the Very Large Telescope (NACO/VLT) on the night of 14 August 2019 in NGS mode with the Ks filter. We took nine frames with an integration time of 17 s each, and dithered between each frame. We performed a standard reduction using a custom IDL pipeline: we subtracted flats and constructed a sky background from the dithered science frames, aligned and co-added the images, and then injected fake companions to determine a 5σ detection threshold as a function of radius. We obtained a contrast of 5.6 magnitudes at 1″, and no companions were detected. The contrast curve is shown in Extended Data Fig. 4.

SOAR. We searched for nearby sources to TOI-849 with SOAR (Southern Astrophysical Research) speckle imaging[43,44] on 12 August 2019 UT, observing in a similar visible bandpass as TESS. We detected no nearby sources within 3″ of TOI-849. The 5σ detection sensitivity and the speckle auto-correlation function from the SOAR observation are plotted in Extended Data Fig. 4.

AstraLux. We obtained a high-spatial-resolution image of TOI-849 with the AstraLux camera[45], which is installed at the 2.2-m telescope of Calar Alto Observatory (Almería, Spain), using the 'lucky imaging' technique[46]. We obtained 24,400 images in the SDSSz band with 20 ms exposure time, well below the coherence time. The CCD was windowed to match 6″ × 6″. We used the observatory pipeline to perform basic reduction of the images and subsequent selection of the best-quality frames. This was done by measuring their Strehl ratio[47] and selecting only the 10% with the highest value of this parameter (an effective integration time of 48 s). Then, these images were aligned and combined to obtain the final high-spatial-resolution image. We estimated the sensitivity curve of this high-spatial-resolution image[48,49] based on the injection of artificial stars in the image at different angular separations and position angles and by measuring the retrieved stars using the detection algorithms used to look for real companions. No companions were detected in this image within the sensitivity limits. Both the high-resolution image and the contrast curve are shown in Extended Data Fig. 4.

Zorro/Gemini South. TOI-849 was observed on 13 September 2019 UT using the Zorro speckle instrument on the Gemini South telescope. Zorro provides simultaneous speckle imaging in two bands, 562 nm and 832 nm, with output data products including a reconstructed image and with robust limits on companion detections[50]. Extended Data Fig. 4 shows our 562-nm contrast curve from which we find that TOI-849 is a single star with no companion brighter than about 5 magnitudes detected within 1.75″.

Spectroscopic analysis and chemical abundances. The spectroscopic analysis used to derive the effective temperature (T_{eff}), surface gravity (log*g*), microturbulence (ξ_{t}) and metallicity ([Fe/H]) and the respective errors follows previous work[51,52]. Equivalent widths are measured for a list of well defined iron lines. We use the combined HARPS spectrum of TOI-849 and ARES v2 code[53,54] to measure the equivalent widths. In the spectral analysis we look for the ionization and excitation equilibrium. The process makes use of a grid of Kurucz model atmospheres[55] and the radiative-transfer code MOOG[56]. The resulting values are $T_{\mathrm{eff}} = 5{,}329 \pm 48$ K, $\log g = 4.28 \pm 0.09$, $\xi_{\mathrm{t}} = 0.82 \pm 0.08$ and [Fe/H] = 0.20 ± 0.03.

The same tools and models are also used to derive stellar abundances for several chemical elements. For this, we use the classical curve-of-growth analysis method assuming local thermodynamic equilibrium. Although the equivalent widths of the spectral lines are automatically measured with ARES, for elements with only two to three lines available we perform careful visual inspection of the equivalent width measurements. Chemical abundances are derived by closely following past work[57]. The final abundances derived are [Na I/H] = 0.30 ± 0.16, [Mg I/H] = 0.24 ± 0.06, [Al I/H] = 0.30 ± 0.06, [Si I/H] = 0.24 ± 0.08, [Ca I/H] = 0.16 ± 0.07, [Sc II/H] = 0.23 ± 0.09, [Ti I/H] = 0.25 ± 0.09, [Cr I/H] = 0.23 ± 0.07 and [Ni I/H] = 0.28 ± 0.04.

Extended Data Figure 5 shows a comparison of the abundances of TOI-849 with those found in solar-neighbourhood stars[58] of similar atmospheric parameters. In terms of chemical composition, TOI-849 seems to be very similar to the solar-neighbourhood stars, showing a slight enhancement in the iron-peak elements Cr and Ni.

Joint RV and photometric fit. The HARPS RVs, the TESS, NGTS and LCOGT photometry, and the spectral energy distribution (SED) were jointly analysed in a Bayesian framework, using the PASTIS software[11,59]. For the SED, we used the visible magnitudes from the American Association of Variable Star Observers Photometric All-Sky Survey (APASS) and the near-infrared magnitudes from the Two-Micron All-Sky Survey (2MASS) and the Wide-field Infrared Survey Explorer (AllWISE)[60–62]. The RVs were fitted using a Keplerian orbit model and a linear drift. The light curves were modelled with the JKT Eclipsing Binary Orbit Program[63] using an oversampling factor of 180, 12, 6 and 7 for the TESS and the three LCOGT-CTIO light curves, respectively. The NGTS light curves were not oversampled because the integration of the individual data is short with respect to the transit duration[64]. Finally, the SED was modelled using the BT-Settl library of stellar atmosphere models[65]. The system parameters and associated uncertainties were derived using the Markov chain Monte Carlo (MCMC) method implemented in PASTIS. The stellar parameters were computed using the Dartmouth evolution tracks[66] at each step of the chains, accounting for the asterodensity profiling[67]. We also used the PARSEC evolution tracks, with consistent results.

Regarding the priors, we used a normal distribution with median and width from the spectral analysis to obtain the stellar temperature, surface gravity and iron abundance. For the systemic distance to Earth, we used a normal prior centred on the Gaia Data Release 2 value[10], taking into account the distance bias correction[68]. For the orbital period and transit epoch, we used normal priors centred on first-guess values from an independent analysis of the NGTS and TESS light curves alone, to improve the convergence of the MCMCs. For the orbital inclination we used a sine prior and for the eccentricity a truncated normal prior with width 0.083 (ref. [69]). For the other parameters, we used uniform priors with width large enough to not artificially decrease the uncertainties. Initial fits gave an eccentricity of 0.036 ± 0.027 (1σ error; result indistinguishable from zero), so we fixed the eccentricity to zero for the final fitting. A linear drift was included for the HARPS data that was also indistinguishable from zero and did not affect the results. Further testing with a quadratic drift model showed no changes in the fit parameters and was dropped.

We ran 20 MCMCs with 2×10^5 iterations. We checked the convergence with a Kolmogorov–Smirnov test[11,59], removed the burn-in phase and merged the remaining chains. The limb-darkening coefficients were computed using previously computed stellar parameters and tables[70]. Finally, the physical parameters and associated uncertainties were derived from samples from the merged chain. The results for the Dartmouth and PARSEC evolution tracks are shown in Extended Data Tables 1, 2. The fit transit depth implies a joint signal-to-noise ratio of 386 (ref. [59]) for the transit.

As an independent check on the derived stellar parameters, we performed an analysis of the broadband SED together with the Gaia parallax to determine an empirical value of the stellar radius[17,71,72]. We pulled the $B_{\mathrm{T}}V_{\mathrm{T}}$ magnitudes from Tycho-2, the *BV*gri magnitudes from APASS, the JHK_{S} magnitudes from 2MASS, the *W*1–*W*4 magnitudes from WISE and the *G* magnitude from Gaia. Together, the available photometry spans the full stellar SED over the wavelength range 0.4–22 μm. We also checked the GALEX near-ultraviolet flux, which was not used in the fit, because it suggests a modest level of chromospheric activity.

We performed the independent fit using the Kurucz stellar atmosphere models, with priors on T_{eff}, log*g* and [Fe/H] from the spectroscopic values. The remaining free parameter is the extinction (A_V), which we limited to the maximum line-of-sight extinction from known dust maps[73]. The resulting fit has a reduced χ^2 of 4.5 and a best-fit extinction of $A_V = 0.04 \pm 0.03$. Integrating the (unextincted) model SED gives a bolometric flux at Earth of $F_{\mathrm{bol}} = 3.713 \pm 0.086 \times 10^{-10}$ erg s cm^{-2}. Taking F_{bol} and T_{eff} together with the Gaia parallax, adjusted by +0.08 mas to account for a previously reported systematic offset[74], gives the stellar radius as $R = (0.896 \pm 0.020)R_{\odot}$. Finally, estimating the stellar mass from known empirical relations[75], assuming solar metallicity, gives $M = (1.01 \pm 0.08)M_{\odot}$, which, combined with the radius, gives a mean stellar density of $\rho = 1.99 \pm 0.19$ g cm^{-3}. These values are consistent with the stellar parameters found from the PASTIS MCMC chain, so we adopt the PASTIS values for our results.

Interpretation and discussion

Interior-structure characterization. Given the mass and radius of TOI-849b, it is clear that the planet does not represent a larger version of Neptune. This is demonstrated in Fig. 2, which shows the mass–radius relation for a pure-water curve and a planet consisting of 95% water and 5% H–He atmosphere, corresponding to a stellar irradiation of $F/F_{\oplus} = 3{,}000$ (TOI-849b). TOI-849b sits on the pure-water curve and well below the 5% strongly irradiated curve, suggesting that the H–He mass fraction is of the order of only a few per cent, if not negligible. Figure 3 also shows that TOI-849b is relatively isolated in parameter space, suggesting that it is somewhat unique and could have been subjected to an unusually aggressive removal of the primordial H–He envelope.

We explore layered-structure models containing variable fractions of the H–He envelope. Typical available models are not suited to this planet owing to the high pressures in the interior, requiring exotic equations of state. Further, for planets this massive, the interior layers are probably not as distinct as for smaller planets, with composition gradients more likely[12]. Rather than build a full model of the interior, which would not be valid for the reasons stated, we consider some illuminating limiting cases.

We model the planetary interior of TOI-849b assuming a pure iron core, a silicate mantle, a pure water layer, and a H–He atmosphere. We build a structure model based on previous work[13] except for the iron core, for which we use an updated equation of state[76]. For the silicate mantle, the equilibrium mineralogy and density are computed as a function of pressure, temperature and bulk composition by minimizing the Gibbs free energy[77]. For the water, we use a quotidian equation of state[78] for low pressures and a previously tabulated equation of state[79] for pressures above 44.3 GPa. For H–He, we assume a proto-solar composition[80]. We then solve the standard structure equations.

We then estimate the possible range of the H–He mass fraction in TOI-849b that fits the derived mass and radius. To estimate the maximum possible mass of an H–He envelope, we assume a planet without water. The core-to-mantle fraction is set by the stellar abundance [Fe/Si] of the host star[81]. The minimum H–He mass fraction is estimated by assuming a large fraction of water of 70% by mass, which corresponds to a water-rich planet. We search for the maximum and minimum H–He mass fractions for a grid of planetary masses and radii covering the observed values and their 2σ error range. It is found that that H–He mass fraction is at minimum $2.9^{+0.8}_{-1.0}\%$ and at maximum $3.9^{+0.8}_{-0.9}\%$, suggesting that the heavy-element mass is greater than $38M_{\oplus}$. It should be noted that our models assume a pure H–He atmosphere, whereas in reality the atmosphere is expected to include heavier elements, as inferred by recent formation models[82]. This is particularly true for planets this massive, where the interior layers are probably not as distinct as for smaller planets. The existence of heavy elements in the H–He atmosphere would lead to compression, and can therefore increase the planetary H–He mass fraction. However, for the case of TOI-849b, the difference is expected to be very moderate because the planet mass is clearly dominated by heavy elements. Previous work calculated the effect of varying atmospheric water content on planetary radii for fixed masses and H–He gas mass fractions[83]. Applying that model to TOI-849b shows that the inferred planet radius is only affected on the level of a few per cent for atmospheric water content ranging from 0 to 70%. As such, we expect the plausible increase in H–He to be small even for high levels of volatile enrichment in the planetary envelope. We can therefore conclude that the mass fraction of H–He is at most a few per cent.

Photoevaporation rate. We explored the X-ray and extreme ultraviolet irradiation of the planet—the wavelengths most relevant for atmospheric mass loss[84]. Archival X-ray data exist for the system only from the ROSAT all-sky survey, where the nearest detected source is an arcminute away, too far to be associated with TOI-849. Instead, we applied known empirical relations linking X-ray emission with age[85], estimating $L_X/L_{\text{bol}} = 7.5 \times 10^{-7}$ at the current age, where L_X is the X-ray luminosity and L_{bol} the bolometric luminosity. This figure implies an X-ray flux at Earth of 3.0×10^{-16} erg s^{-1} cm^{-2}, much too faint to be visible with XMM-Newton or Chandra. We extrapolated our X-ray estimate to the unobservable extreme ultraviolet band using previously derived relations[86,87].

To estimate mass-loss rates, we applied both the energy-limited approach[88,89], and a method based on interpolating and approximating to hydrodynamical simulations[90,91]. The latter yields a loss rate of 1.8×10^{11} g s^{-1}, more than an order of magnitude larger than the former when assuming a canonical efficiency of 15%. By integrating over the planet's extreme ultraviolet history, and starting at a Jupiter mass and radius, we estimate total lifetime losses of 4.0% and 0.81% of the planet's mass using the energy-limited and Kubyshkina methods, respectively. Although these calculations have the limitation of assuming a constant radius across the lifetime, these losses are not enough to evolve the planet to one slightly smaller than Neptune, so we can be sure that the planet did not start as a Jupiter-like giant if its evolution has been solely through photoevaporation.

An intermediate starting point is the planet HD149026b[3], a giant planet with mass $(121 \pm 19)M_{\oplus}$ and radius $(8.3 \pm 0.2)R_{\oplus}$ (ref. [17]). For this planet, we estimate total lifetime losses of 11.42% and 100% of the planet's mass using the energy-limited and Kubyshkina methods, respectively. These are likely to be overestimates owing to the constant-radius assumption, which clearly becomes flawed after mass loss representing a large fraction of the planet's mass. As such, finding the limits of photoevaporation in creating a planet like TOI-849b requires detailed models beyond the scope of this paper.

Co-orbital bodies and exomoons. The anomalously large density found for planet TOI-849b allows us to explore alternative scenarios for the origin of this signal. One of the most relevant ones is the co-orbital case. Although these configurations have not yet been confirmed in any extrasolar systems despite several efforts[92–96], some candidates have arisen from studies such as Kepler-91[97] or the recent TOI-178[98]. Indeed, an additional planet in the system with the same orbital period as that of TOI-849b but not transiting owing to a mutual inclination between their orbits (or that is too small to be detected by TESS) could explain the large mass measured for such a small planet radius.

We here explore the scenario in which two planetary-mass bodies share the same orbital period in a 1:1 mean-motion resonance configuration. In such a case, the mass that we measure in the joint fit would be distributed in two planetary-mass objects. Such configurations are allowed by dynamical stability studies, which demonstrate that the only condition for the stability of co-orbital configurations is that the total mass of the planet plus its co-orbital companion must be smaller than 3.8% of the mass of the star[99]. Regardless of the formation process, and given the mass of the star and the estimated mass of TOI-849b, the co-orbital scenario would be stable for any planetary mass of the accompanying body.

To test this hypothesis, we apply a recently derived procedure to analyse the RV of the star using a new RV equation including two Keplerian components[95,96,100]. The new equation can be simplified so that only one extra parameter, α, is included[100]. This parameter depends on the trojan-to-planet mass ratio, so that if positive (negative), a trojan candidate might be in L_5 (L_4). For this analysis, we first assume a circular orbit, thus having five parameters: namely, the RV semi-amplitude K_{coorb}, the orbital period, the main-planet time of conjunction $T_{0,\text{b}}$, the systemic velocity γ and the alpha parameter α. We use Gaussian priors on the orbital period and time of conjunction with the parameters derived from the one-planet analysis (see Extended Data Table 1) and uniform priors for the alpha parameter $\mathcal{U}(-1, 1)$ km s^{-1} and systemic velocity $\mathcal{U}(9.1, 9.5)$ km s^{-1}. We also include a jitter term and a slope.

We use *emcee*[101] with 50 walkers and 5,000 steps per walker to explore the parameter space. We use the last half of each chain to compute the final posterior distributions. For the key parameter α, we obtain $\alpha = -0.092^{+0.060}_{-0.064}$. This value is 1.5σ away from zero and hence compatible with it within a 95% confidence level. The posterior distribution allows us to discard co-orbitals more massive than $8M_\oplus$ at the 95% confidence level assuming a mean resonant angle ζ, where $\zeta = \lambda_1 - \lambda_2$ and λ_i is the mean longitude of each of the two co-orbitals, of 60°. In practice, this assumes that the trojan planet would have been located exactly at the Lagrangian point during the timespan of the observations. In such a case, the transiting planet would have a mass of $31M_\oplus$, still uniquely high for its radius. A particular arrangement of trojan planets whereby equal-mass trojans are present in both the L_4 and L_5 Lagrangian points could in principle mimic the observed HARPS data. Such a scenario is observationally indistinguishable from the single-planet model while being notably more complex, and we reject it on that basis.

A related hypothesis is that of a 'double planet' or moon with non-negligible mass. In such a scenario, there is no distinguishable effect on the RVs and hence the apparent large mass would be split over additional bodies. We estimate the minimum stable satellite density by considering where the Hill radius and Roche limit of the planet overlap for TOI-849b[102]. Equation (5) of ref. [102] gives a minimum stable satellite density of 38 g cm^{-3}, much denser than pure iron. As such, we conclude that physically realistic exomoons are unstable around TOI-849b and this hypothesis can be discarded.

Planet population synthesis models. We explored possible formation channels for such dense Neptune-sized planets using the Bern Generation 3 Model of Planetary Formation and Evolution, which is an update on the currently published version[26]. The main changes in the model are reflected in the following description. The model self-consistently evolves a one-dimensional gas disk, the dynamical state of the solids, the accretion of solids and gas by the protoplanets, their interiors, and their dynamical evolution by gravitational interactions and gas-driven migration.

For the gas disk, the model computes a one-dimensional radial profile that is evolving viscously[103], with the macroscopic viscosity given by the standard α parameterization[104]. The vertical structure is now computed using a vertically integrated approach[105] that includes the effect of stellar irradiation[106]. Stellar parameters are retrieved from known evolution tracks[107]. We include additional sink terms for the accretion by the planets, as well as both internal[108] and external[109] photoevaporation.

The model assumes that planetesimals accrete in the oligarchic regime[110–112], and their capture cross-section is computed consistently with the envelope structure[113]. The internal structure equations[114] are solved for the gas envelope. In the initial (or 'attached') phase, the envelope is in equilibrium with the surrounding disk, and the internal structure is used to determine the gas mass. Gas accretion is governed by the ability of the planet to radiate away the gravitational energy released from the accretion of both solids and gas[115,116]. When the accretion rate exceeds the supply from the disk, the envelope is no longer in equilibrium with the disk and contracts[117]. In this 'detached' phase, the internal structure is used to retrieve the planet's radius and luminosity.

Dynamical interactions between the planets are simulated by means of the Mercury N-body integrator[118]. After a giant impact, an additional luminosity is included[119] to determine whether the gas envelope is ejected. Gas-driven type I migration is computed in line with past work[120], accounting for how local thermodynamic effects in the disk[121] and planet eccentricities and inclinations[122] affect the corotation torques. Type II migration and the switch between the two migration regimes are computed in line with past work[123]. Torques and damping are included in the N-body calculation by means of additional forces.

The formation stage lasts for 20 Myr. The model then transitions into the evolution stage, in which the planets are followed individually for up to 10 Gyr. This stage includes thermodynamical evolution of the envelope, atmospheric escape[124,125] and tidal migration[126] with a fixed stellar dissipation parameter of $Q_\star = 10^6$.

To obtain a synthetic population, we update a previously published procedure[127]. We use the literature disk mass distribution[128], and the characteristic radius, which determines the radial distribution of the gas, is obtained following a known relationship[129]. The location of the inner edge of the disk has a log-normal distribution in period with a mean of 4.7 d (ref. [130]). The dust-to-gas ratio is obtained from the observed stellar [Fe/H] (ref. [127]), but using the primordial solar metallicity as a reference[131]. The initial surface density profile of solids has a steeper slope than that of the gas[132], leading to a higher concentration in the inner region. In each disk, 20 lunar-mass ($10^{-2}M_\oplus$) planetary embryos are emplaced at the beginning. Their initial positions are randomly selected between the inner edge of the disk up to 40 AU, with a uniform probability in the logarithm of the semi-major axis.

In those models, which were run before the discovery of TOI-849b, we found three planets that exhibit similar mass, radius and eccentricity to TOI-849b out of a total sample of 1,000. These planets have masses between $20M_\oplus$ and $50M_\oplus$ and have an ice content of 20–30% by mass, but no H/He. They started as embryos outside the ice line and migrated steadily to a position close to the inner edge of the disk. The removal of the primordial H/He is due to one or two giant impacts that took place at the end of the migration, which means that the planets were unable to accrete a second H/He envelope. For one of the three planets only a single impact is seen, whereas two impacts occur in the others. In all cases, only a single impact is needed to remove the envelope. To place this in context, 70% of close-in Neptunes in the simulations, defined as having a semi-major axis <0.04 AU, had at least one impact with a body of mass $>1M_\oplus$ during their formation. As such, impacts are not particularly rare, but the timing of the impact at the end of the migration phase is what prevents reaccretion and leads to a permanently lost envelope.

Owing to the high equilibrium temperature, it is likely that the remaining ices evaporate to form a secondary atmosphere consisting of water and possibly other volatiles like CO and CO_2. Such an envelope leads to radii comparable to that of the discovered planet. From the modelling point of view, the population synthesis models prefer planets with small envelopes consisting entirely of ices. The evolution tracks of the three considered model planets are shown in Extended Data Fig. 6.

Although no model planets similar to TOI-849b were found from other formation pathways, this should not be taken as evidence against other hypotheses, such as gap opening limiting the accretion, or tidal disruption. The Bern models do not include gap opening in the disk as a limiting factor in gas accretion and use simplified assumptions for tidal interactions[126] that do not include high-eccentricity migration.

Tidally induced thermalization events. The high bulk density of TOI-849b (5.2 g cm^{-3}) relative to that of Neptune (1.6 g cm^{-3}) suggests that the planet (with a radius equal to 90% of Neptune's) might currently represent the core of a previously giant planet. For this scenario to be viable, the planet needs to originate from a gas giant and to have expelled mass, possibly during orbit shrinkage and circularization. This evolutionary pathway may occur as a result of chaotic tides[133–135], where the planet's internal f-modes were excited after the planet was gravitationally scattered onto a highly eccentric orbit. Energy buildup in the modes could have then led to thermalization events, potentially ejecting atmospheric layers[136,137]. After the resulting core left the chaotic regime, subsequent orbital evolution over the ~9-Gyr main-sequence lifetime of the parent star may have proceeded with weakly dissipative equilibrium tides, leading to the current orbit. In this scenario, the planet may have expelled 1–2 orders of magnitude more mass than its current value.

Accumulation of the internal mode energy leads to thermalization events, which subsequently deposit energy into the planet's interior and reset the mode amplitude. Possible results of the thermalization events include inflation, mass ejection or both; TOI-849b could have experienced such events and still retained some or all of its atmosphere. Although their trigger and consequences remain largely unknown, previous work has assumed that such events occur when the accumulated mode energy equals 10% of the planet's binding energy[136]

$$E_{\mathrm{bind}} \approx \frac{GM_{\mathrm{p}}^2}{R_{\mathrm{p}}}, \qquad (1)$$

where M_{p} and R_{p} are the mass and radius of the planet, respectively, and G is the gravitational constant. That work[136] also demonstrated that the changes in orbital evolution resulting from the thermalization events are largely independent of this choice of 10%. With this selection, it has been illustrated that the number of thermalization events that a planet experiences is positively correlated with an increasing puffiness of the planet and a decreasing orbital pericentre[137]. It has been shown that even a dense gas giant with a pericentre of about $1.5R_{\odot}$ would experience at least one thermalization event, albeit with a smaller-mass central star. TOI-849b, which currently resides at a distance of about $3R_{\odot}$, previously would have harboured a pericentre that is just half of that value if angular momentum was conserved as its eccentricity decreased from almost unity to zero, under the high-eccentricity circularization scenario.

Atmospheric follow-up observations. Future observations of TOI-849b may attempt to identify its atmospheric composition. TOI-849b represents a new class of dense, high-mass planet and its atmosphere will provide a counterpoint to other planets of different type, as well as potentially allow the characterization of a non-H_2-rich atmosphere. Given the high equilibrium temperature of the planet, and hence the potential for evaporation of volatiles to form a secondary atmosphere, such observations may be able to detect core material in the atmosphere, and regardless will help to place TOI-849b in context against other Neptune-sized planets, other planets with or without high irradiation, the few planets inside the Neptunian desert and the bulk composition of the star. Such comparisons are the goal of the European Space Agency's Ariel mission[138], although the magnitude of TOI-849 will arguably require next-generation telescopes for atmospheric observations, such as the James Webb Space Telescope or the European Extremely Large Telescope.

Data availability

TESS data are publicly available at MAST (https://archive.stsci.edu/missions-and-data/transiting-exoplanet-survey-satellite-tess). The HARPS data used in this study are available within the paper or the Supplementary Information files and were collected under ESO programme ID 1102.C-0249. The NGTS (Data Tags 19249 and 19250), LCOGT (Data Tags 5106 and 5386) and specific detrended TESS light curve (Data Tag 19248) used in this work are available via the Exofop-TESS archive (https://exofop.ipac.caltech.edu/tess/).

Code availability

The PASTIS code has been published previously[11,59]. The latest version of the ARES code (ARES v2) is available at http://www.astro.up.pt/~sousasag/ares.

32. Jenkins, J. M. et al. The TESS science processing operations center. In *Proc. SPIE Software and Cyberinfrastructure for Astronomy IV Series*, Vol. 9913, 99133 (SPIE, 2016).
33. Huang, X. et al. A quick look into the first discoveries of TESS. In *American Astronomical Society Meeting Abstracts* Vol. 233, 209.08 (American Astronomical Society, 2019).
34. Vanderburg, A. & Johnson, J. A. A technique for extracting highly precise photometry for the two-wheeled Kepler mission. *Publ. Astron. Soc. Pacific* **126**, 948 (2014).
35. Armstrong, D. J. et al. K2 variable catalogue: variable stars and eclipsing binaries in K2 campaigns 1 and 0. *Astron. Astrophys.* **579**, A19 (2015).
36. McCormac, J. et al. DONUTS: a science frame autoguiding algorithm with sub-pixel precision, capable of guiding on defocused stars. *Publ. Astron. Soc. Pacific* **125**, 548 (2013).
37. Mayor, M. et al. Setting new standards with HARPS. *Messenger* **114**, 20–24 (2003).
38. Baranne, A. et al. ELODIE: a spectrograph for accurate radial velocity measurements. *Astron. Astrophys. Suppl. Ser*. **119**, 373–390 (1996).
39. Pepe, F. et al. HARPS: ESO's coming planet searcher. Chasing exoplanets with the La Silla 3.6-m telescope. *Messenger* **110**, 9–14 (2002).
40. Boisse, I. et al. Disentangling between stellar activity and planetary signals. *Astron. Astrophys*. **528**, A4 (2011).
41. Jensen, E. Tapir: A web interface for transit/eclipse observability. *Astrophysics Source Code Library* **1306**.007 (2013).
42. Collins, K. A., Kielkopf, J. F., Stassun, K. G. & Hessman, F. V. AstroImageJ: image processing and photometric extraction for ultra-precise astronomical light curves. *Astron. J*. **153**, 77 (2017).
43. Tokovinin, A. Ten years of speckle interferometry at SOAR. *Publ. Astron. Soc. Pacific* **130**, 035002 (2018).
44. Ziegler, C. et al. SOAR TESS Survey. I: sculpting of TESS planetary systems by stellar companions. *Astron. J*. **159**, 19 (2019).
45. Hormuth, F., Brandner, W., Hippler, S. & Henning, T. AstraLux – the Calar Alto 2.2-m telescope LuckyImaging camera. *J. Phys. Conf. Ser*. **131**, 012051 (2008).
46. Fried, D. L. Probability of getting a lucky short-exposure image through turbulence. *J. Opt. Soc. Am*. **68**, 1651–1658 (1978).
47. Strehl, K. Über die Bildschärfe der Fernrohre. *Astron. Nachr*. **158**, 89–90 (1902).
48. Lillo-Box, J., Barrado, D. & Bouy, H. Multiplicity in transiting planet-host stars. A lucky imaging study of Kepler candidates. *Astron. Astrophys*. **546**, A10 (2012).
49. Lillo-Box, J., Barrado, D. & Bouy, H. High-resolution imaging of Kepler planet host candidates. A comprehensive comparison of different techniques. *Astron. Astrophys*. **566**, A103 (2014).
50. Howell, S. B., Everett, M. E., Sherry, W., Horch, E. & Ciardi, D. R. Speckle camera observations for the NASA Kepler Mission Follow-up Program. *Astron. J*.**142**, 19 (2011).
51. Sousa, S. G. et al. A new procedure for defining a homogenous line-list for solar-type stars. *Astron. Astrophys*. **561**, A21 (2014).
52. Santos, N. C. et al. SWEET-Cat: a catalogue of parameters for Stars With ExoplanETs. I. New atmospheric parameters and masses for 48 stars with planets. *Astron. Astrophys*. **556**, A150 (2013).
53. Sousa, S. G., Santos, N. C., Israelian, G., Mayor, M. & Monteiro, M. J. P. F. G. A new code for automatic determination of equivalent widths: Automatic Routine for line Equivalent widths in stellarSpectra (ARES). *Astron. Astrophys*. **469**, 783–791 (2007).
54. Sousa, S. G., Santos, N. C., Adibekyan, V., Delgado-Mena, E. & Israelian, G. ARES v2: new features and improved performance. *Astron. Astrophys*. **577**, A67 (2015).
55. Kurucz, R. L. *SYNTHE Spectrum Synthesis Programs and Line Data CD-ROM* (Smithsonian Astrophysical Observatory, 1993).
56. Sneden, C. A. *Carbon and Nitrogen Abundances in Metal-Poor Stars. PhD thesis*, Univ. of Texas at Austin (1973).
57. Adibekyan, V. Z. et al. Chemical abundances and kinematics of 257 G-, K-type field giants. Setting a base for further analysis of giant-planet properties orbiting evolved stars. *Mon. Not. R. Astron. Soc*. **450**, 1900–1915 (2015).
58. Adibekyan, V. Z. et al. Chemical abundances of 1111 FGK stars from the HARPS GTO planet search program. Galactic stellar populations and planets. *Astron. Astrophys*. **545**, A32 (2012).
59. Díaz, R. F. et al. PASTIS: Bayesian extrasolar planet validation – I. General framework, models, and performance. *Mon. Not. R. Astron. Soc*. **441**, 983–1004 (2014).
60. Henden, A. A., Levine, S., Terrell, D. & Welch, D. L. APASS – the latest data release. In *American Astronomical Society Meeting Abstracts* Vol. 225, 336.16 (2015).
61. Munari, U. et al. APASS Landolt-Sloan BVgri photometry of RAVE Stars. I. Data, effective temperatures, and reddenings. *Astron. J*. **148**, 81 (2014).
62. Cutri, R. M. & et al. VizieR Online Data Catalog: AllWISE Data Release (Cutri+ 2013) 2328 http://vizier.u-strasbg.fr/viz-bin/VizieR?-source=II/328 (2014).
63. Southworth, J. Homogeneous studies of transiting extrasolar planets - I. Light-curve analyses. *Mon. Not. R. Astron. Soc*. **386**, 1644–1666 (2008).
64. Kipping, D. M. Binning is sinning: morphological light-curve distortions due to finite integration time. *Mon. Not. R. Astron. Soc*. **408**, 1758–1769 (2010).
65. Allard, F., Homeier, D. & Freytag, B. Models of very-low-mass stars, brown dwarfs and exoplanets. *Philos. Trans. Royal Soc. A* **370**, 2765–2777 (2012).
66. Dotter, A. et al. The Dartmouth stellar evolution database. Astrophys. *J. Suppl. Ser*. **178**, 89–101 (2008).
67. Kipping, D. M. Characterizing distant worlds with asterodensity profiling. *Mon. Not. R. Astron. Soc*. **440**, 2164–2184 (2014).
68. Schönrich, R., McMillan, P. & Eyer, L. Distances and parallax bias in Gaia DR2. *Mon. Not. R. Astron. Soc*. **487**, 3568–3580 (2019).
69. Van Eylen, V. et al. The orbital eccentricity of small planet systems. *Astron. J*. **157**, 61 (2019).
70. Claret, A. & Bloemen, S. Gravity and limb-darkening coefficients for the Kepler, CoRoT, Spitzer, uvby, UBVRIJHK, and Sloan photometric systems. *Astron. Astrophys*. **529**, A75 (2011).
71. Stassun, K. G. & Torres, G. Eclipsing binaries as benchmarks for trigonometric parallaxes in the Gaia era. *Astron. J*. **152**, 180 (2016).
72. Stassun, K. G., Corsaro, E., Pepper, J. A. & Gaudi, B. S. Empirical accurate masses and radii of single stars with TESS and Gaia. *Astron. J*. **155**, 22 (2018).
73. Schlegel, D. J., Finkbeiner, D. P. & Davis, M. Maps of dust infrared emission for use in estimation of reddening and cosmic microwave background radiation foregrounds. *Astrophys. J*. **500**, 525–553 (1998).
74. Stassun, K. G. & Torres, G. Evidence for a systematic offset of −80 µas in the Gaia DR2 parallaxes. *Astrophys. J*. **862**, 61 (2018).
75. Torres, G., Andersen, J. & Giménez, A. Accurate masses and radii of normal stars: modern results and applications. *Astron. Astrophys*. **18**, 67–126 (2010).
76. Hakim, K. et al. A new ab initio equation of state of hcp-Fe and its implication on the interior structure and mass-radius relations of rocky super-Earths. *Icarus* **313**, 61–78 (2018).
77. Connolly, J. A. D. The geodynamic equation of state: what and how. *Geochem. Geophys. Geosyst*. **10**, Q10014 (2009).

78. Vazan, A., Kovetz, A., Podolak, M. & Helled, R. The effect of composition on the evolution of giant and intermediate-mass planets. *Mon. Not. R. Astron. Soc*. **434**, 3283–3292 (2013).
79. Seager, S., Kuchner, M., Hier-Majumder, C. A. & Militzer, B. Mass–radius relationships for solid exoplanets. *Astrophys. J*. **669**, 1279–1297 (2007).
80. Saumon, D., Chabrier, G. & van Horn, H. M. An equation of state for low-mass stars and giant planets. *Astrophys. J. S* **99**, 713 (1995).
81. Dorn, C. et al. Can we constrain the interior structure of rocky exoplanets from mass and radius measurements? *Astron. Astrophys*. **577**, A83 (2015).
82. Lozovsky, M., Helled, R., Rosenberg, E. D. & Bodenheimer, P. Jupiter's formation and its primordial internal structure. *Astrophys. J*. **836**, 227 (2017).
83. Lozovsky, M., Helled, R., Dorn, C. & Venturini, J. Threshold radii of volatile-rich planets. *Astrophys. J*. **866**, 49 (2018).
84. Lammer, H. et al. Atmospheric loss of exoplanets resulting from stellar X-Ray and extreme-ultraviolet heating. *Astrophys. J. Lett*. **598**, 121–124 (2003).
85. Jackson, A. P., Davis, T. A. & Wheatley, P. J. The coronal X-ray-age relation and its implications for the evaporation of exoplanets. *Mon. Not. R. Astron. Soc*. **422**, 2024–2043 (2012).
86. King, G. W. et al. The XUV environments of exoplanets from Jupiter-size to super-Earth. *Mon. Not. R. Astron. Soc*. **478**, 1193–1208 (2018).
87. Chadney, J. M., Galand, M., Unruh, Y. C., Koskinen, T. T. & Sanz-Forcada, J. XUV-driven mass lossfrom extrasolar giant planets orbiting active stars. *Icarus* **250**, 357–367 (2015).
88. Watson, A. J., Donahue, T. M. & Walker, J. C. G. The dynamics of a rapidly escaping atmosphere: applications to the evolution of Earth and Venus. *Icarus* **48**, 150–166 (1981).
89. Erkaev, N. V. et al. Roche lobe effects on the atmospheric loss from "Hot Jupiters". *Astron. Astrophys*. **472**, 329–334 (2007).
90. Kubyshkina, D. et al. Grid of upper atmosphere models for 1–40 $M_{\oplus}$ planets: application to CoRoT-7b and HD 219134b,c. *Astron. Astrophys*. **619**, A151 (2018).
91. Kubyshkina, D. et al. Overcoming the limitations of the energy-limited approximation for planet atmospheric escape. *Astrophys. J. Lett*. **866**, 18 (2018).
92. Madhusudhan, N. & Winn, J. N. Empirical constraints on trojan companions and orbital eccentricities in 25 transiting exoplanetary systems. *Astrophys. J*. **693**, 784–793 (2009).
93. Ford, E. B. & Gaudi, B. S. Observational constraints on trojans of transiting extrasolar planets. *Astrophys. J. Lett*. **652**, 137–140 (2006).
94. Janson, A. Systematic search for trojan planets in the Kepler data. *Astrophys. J*. **774**, 156 (2013).
95. Lillo-Box, J. et al. The TROY project: searching for co-orbital bodies to known planets. I. Project goals and first results from archival radial velocity. *Astron. Astrophys*. **609**, A96 (2018).
96. Lillo-Box, J. et al. The TROY project. II. Multi-technique constraints on exotrojans in nine planetary systems. *Astron. Astrophys*. **618**, A42 (2018).
97. Lillo-Box, J. et al. Kepler-91b: a planet at the end of its life. Planet and giant host star properties via lightcurve variations. *Astron. Astrophys*. **562**, A109 (2014).
98. Leleu, A. et al. Co-orbital exoplanets from close period candidates: the TOI-178 case. *Astron. Astrophys*. **624** A46 (2019).
99. Laughlin, G. & Chambers, J. E. Extrasolar trojans: the viability and detectability of planets in the 1:1 resonance. *Astron. J*. **124**, 592–600 (2002).
100. Leleu, A., Robutel, P., Correia, A. C. M. & Lillo-Box, J. Detection of co-orbital planets by combining transit and radial-velocity measurements. *Astron. Astrophys*. **599**, L7 (2017).
101. Foreman-Mackey, D., Hogg, D. W., Lang, D. & Goodman, J. emcee: the MCMC hammer. *Publ. Astron. Soc. Pacif*. **125**, 306–312 (2013).
102. Kane, S. R. Worlds without moons: exomoon constraints for compact planetary systems. *Astrophys. J. Lett*. **839**, L19 (2017).
103. Lynden-Bell, D. & Pringle, J. E. The evolution of viscous discs and the origin of the nebular variables. *Mon. Not. R. Astron. Soc*. **168**, 603–637 (1974).
104. Shakura, N. I. & Sunyaev, R. A. Black holes in binary systems. Observational appearance. *Astron. Astrophys*. **500**, 33–51 (1973).
105. Nakamoto, T. & Nakagawa, Y. Formation, early evolution, and gravitational stability of protoplanetary disks. *Astrophys. J*. **421**, 640 (1994).
106. Hueso, R. & Guillot, T. Evolution of protoplanetary disks: constraints from DM Tauri and GM Aurigae. *Astron. Astrophys*. **442**, 703–725 (2005).
107. Baraffe, I., Homeier, D., Allard, F. & Chabrier, G. New evolutionary models for pre-main sequence and main sequence low-mass stars down to the hydrogen-burning limit. *Astron. Astrophys*. **577**, A42 (2015).
108. Clarke, C. J., Gendrin, A. & Sotomayor, M. The dispersal of circumstellar discs: the role of the ultraviolet switch. *Mon. Not. R. Astron. Soc*. **328**, 485–491 (2001).
109. Matsuyama, I., Johnstone, D. & Hartmann, L. Viscous diffusion and photoevaporation of stellar disks. *Astrophys. J*. **582**, 893–904 (2003).
110. Ida, S. & Makino, J. Scattering of planetesimals by a protoplanet: slowing down of runaway growth. *Icarus* **106**, 210–227 (1993).
111. Ohtsuki, K., Stewart, G. R. & Ida, S. Evolution of planetesimal velocities based on three-body orbital integrations and growth of protoplanets. *Icarus* **155**, 436–453 (2002).
112. Thommes, E. W., Duncan, M. J. & Levison, H. F. Oligarchic growth of giant planets. *Icarus* **161**, 431–455 (2003).
113. Inaba, S. & Ikoma, M. Enhanced collisional growth of a protoplanet that has an atmosphere. *Astron. Astrophys*. **410**, 711–723 (2003).
114. Bodenheimer, P. & Pollack, J. B. Calculations of the accretion and evolution of giant planets: the effects of solid cores. *Icarus* **67**, 391–408 (1986).
115. Pollack, J. B. et al. Formation of the giant planets by concurrent accretion of solids and gas. *Icarus* **124**, 62–85 (1996).
116. Lee, E. J. & Chiang, E. To cool is to accrete: analytic scalings for nebular accretion of planetary atmospheres. *Astrophys. J*. **811**, 41 (2015).
117. Bodenheimer, P., Hubickyj, O. & Lissauer, J. J. Models of the in situ formation of detected extrasolar giant planets. *Icarus* **143**, 2–14 (2000).
118. Chambers, J. E. A hybrid symplectic integrator that permits close encounters between massive bodies. *Mon. Not. R. Astron. Soc*. **304**, 793–799 (1999).
119. Broeg, C. H. & Benz, W. Giant planet formation: episodic impacts versus gradual core growth. *Astron. Astrophys*. **538**, A90 (2012).
120. Coleman, G. A. L. & Nelson, R. P. On the formation of planetary systems via oligarchic growth in thermally evolving viscous discs. *Mon. Not. R. Astron. Soc*. **445**, 479–499 (2014).
121. Paardekooper, S. J., Baruteau, C. & Kley, W. A torque formula for non-isothermal Type I planetary migration - II. Effects of diffusion. *Mon. Not. R. Astron. Soc*. **410**, 293–303 (2011).
122. Fendyke, S. M. & Nelson, R. P. On the corotation torque for low-mass eccentric planets. *Mon. Not. R. Astron. Soc*. **437**, 96–107 (2014).
123. Dittkrist, K. M., Mordasini, C., Klahr, H., Alibert, Y. & Henning, T. Impacts of planet migration models on planetary populations. Effects of saturation, cooling and stellar irradiation. *Astron. Astrophys*. **567**, A121 (2014).
124. Jin, S. et al. Planetary population synthesis coupled with atmospheric escape: a statistical view of evaporation. *Astrophys. J*. **795**, 65 (2014).
125. Jin, S. & Mordasini, C. Compositional imprints in density-distance-time: a rocky composition for close-in low-mass exoplanets from the location of the valley of evaporation. *Astrophys. J*. **853**, 163 (2018).
126. Benítez-Llambay, P., Masset, F. & Beaugé, C. The mass-period distribution of close-in exoplanets. *Astron. Astrophys*. **528**, A2 (2011).
127. Mordasini, C., Alibert, Y. & Benz, W. Extrasolar planet population synthesis. I. Method, formation tracks, and mass-distance distribution. *Astron. Astrophys*. **501**, 1139–1160 (2009).
128. Tychoniec, L. et al. The VLA Nascent Disk and Multiplicity Survey of Perseus Protostars (VANDAM). IV. Free–free emission from protostars: links to infrared properties, outflow tracers, and protostellar disk masses. *Astrophys. J. Suppl. Ser*. **238**, 19 (2018).
129. Andrews, S. M., Wilner, D. J., Hughes, A. M., Qi, C. & Dullemond, C. P. Protoplanetary disk structures in Ophiuchus. II. Extension to fainter sources. *Astrophys. J*. **723**, 1241–1254 (2010).
130. Venuti, L. et al. CSI 2264: investigating rotation and its connection with disk accretion in the young open cluster NGC 2264. *Astron. Astrophys*. **599**, A23 (2017).
131. Lodders, K. Solar system abundances and condensation temperatures of the elements. *Astrophys. J*. **591**, 1220–1247 (2003).
132. Ansdell, M. et al. ALMA survey of Lupus protoplanetary disks. II. Gas disk radii. *Astrophys. J*. **859**, 21 (2018).
133. Ivanov, P. B. & Papaloizou, J. C. B. On the tidal interaction of massive extrasolar planets on highly eccentric orbits. *Mon. Not. R. Astron. Soc*. **347**, 437–453 (2004).
134. Vick, M. & Lai, D. Dynamical tides in highly eccentric binaries: chaos, dissipation, and quasi-steady state. *Mon. Not. R. Astron. Soc*. **476**, 482–495 (2018).
135. Wu, Y. Diffusive tidal evolution for migrating hot Jupiters. *Astron. J*. **155**, 118 (2018).
136. Vick, M., Lai, D. & Anderson, K. R. Chaotic tides in migrating gas giants: forming hot and transient warm Jupiters via Lidov–Kozai migration. *Mon. Not. R. Astron. Soc*. **484**, 5645–5668 (2019).
137. Veras, D. & Fuller, J. Tidal circularization of gaseous planets orbiting white dwarfs. *Mon. Not. R. Astron. Soc*. **489**, 2941–2953 (2019).
138. Tinetti, G. et al. A chemical survey of exoplanets with ARIEL. *Exp. Astron*. **46**, 135–209 (2018).
139. Stassun, K. G. et al. The revised TESS input catalog and candidate target list. *Astron. J*. **158**, 138 (2019).
140. Skrutskie, M. F. et al. The Two Micron All Sky Survey (2MASS). *Astron. J*. **131**, 1163–1183 (2006).

Acknowledgements This paper includes data collected by the TESS missions, which are publicly available from MAST. Funding for the TESS mission is provided by NASA's Science Mission directorate. We acknowledge the use of public TESS Alert data from pipelines at the TESS Science Office and at the TESS Science Processing Operations Center. Resources supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Advanced Supercomputing (NAS) Division at Ames Research Center for the production of the SPOC data products. This research made use of the Exoplanet Follow-up Observation Program website and the NASA Exoplanet Archive, which are operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program. This work makes use of observations from the LCOGT network and is based in part on observations collected at the European Organisation for Astronomical Research in the Southern Hemisphere under ESO programme IDs 1102.C-0249 and P103.C-0449. Some of the observations presented in the paper used the High-Resolution Imaging instrument Zorro at Gemini South (programme ID GS-2019B-Q-111). Zorro was funded by the NASA Exoplanet Exploration Program and built at the NASA Ames Research Center by S.B.H., N. Scott, E. P. Horch and E. Quigley. D.J.A., D.V. and S.L.C. acknowledge support from the STFC via Ernest Rutherford Fellowships ST/R00384X/1, ST/P003850/1 and ST/R003726/1, respectively. M.B. and S.Gandhi acknowledge support from the STFC research grant ST/S000631/1. G.M.K. is supported by the Royal Society as a Royal Society University Research Fellow. F.M. acknowledges support from a Royal Society Dorothy Hodgkin Fellowship. K.G.S. acknowledges partial support from NASA grant 17-XRP17 2-0024. C.Z. is supported by a Dunlap Fellowship at the Dunlap Institute for Astronomy and Astrophysics, funded through an endowment established by the Dunlap family and the University of Toronto. A.W.M. was supported by NASA grant 80NSSC19K0097 to the University of North Carolina at Chapel Hill. D.J.A.B. acknowledges support from the UK Space Agency. C.X.H. and M.N.G. acknowledge support from the Juan Carlos Torres Fellowship. This work was financed by FEDER (Fundo Europeu de Desenvolvimento Regional) funds through the COMPETE 2020 Operational Programme for Competitiveness and Internationalisation (POCI) and by Portuguese funds through FCT (Fundação para a Ciência e a Tecnologia) in the framework of projects UID/FIS/04434/2019; PTDC/FIS-AST/32113/2017 and POCI-01-0145-FEDER-032113; PTDC/FIS-AST/28953/2017 and POCI-01-0145-FEDER-028953. S.G.S., V.A., S.C.C.B. and O.D.S.D. acknowledge support from FCT through Investigador FCT contracts IF/00028/2014/CP1215/CT0002, IF/00650/2015/CP1273/CT0001, IF/01312/2014/CP1215/CT0004 and DL 57/2016/CP1364/CT0004. S.H. acknowledges support from fellowships PD/BD/128119/2016 funded by FCT (Portugal). Work by J.N.W. was partly funded by the Heising-Simons Foundation. C.A.W. acknowledges support from UK Science Technology and Facility Council grant ST/P000312/1. J.L.-B. and D. Barrado are funded by the Spanish State

Research Agency (AEI) Projects ESP2017-87676-C5-1-R and MDM-2017-0737 Unidad de Excelencia María de Maeztu – Centro de Astrobiología (INTA-CSIC). J.S.J. acknowledges funding by Fondecyt through grant 1161218 and partial support from CATA-Basal (PB06, Conicyt). J.I.V. acknowledges support from CONICYT-PFCHA/Doctorado Nacional-21191829, Chile. The French group acknowledges financial support from the French Programme National de Planétologie (PNP, INSU). F.M. acknowledges support from the Royal Society Dorothy Hodgkin Fellowship. C.M. and A.E. acknowledge support from the Swiss National Science Foundation under grant BSSGI0_155816 "PlanetsInTime". Parts of this work have been carried out within the framework of the NCCR PlanetS supported by the Swiss National Science Foundation.

Author contributions D.J.A. is Principal Investigator of the NCORES HARPS programme, which measured the planet's mass, and a member of the NGTS consortium; D.J.A. developed much of the text and main figures and coordinated all contributions. T.A.L. performed the joint PASTIS analysis. V.A., S.G.S. and N.C.S. performed stellar spectral analysis including chemical abundances. R.A.B. and F.M. provided text analysing potential formation scenarios. K.A.C. and E.L.N.J. coordinated the TFOP SG1 photometric follow-up of the system. K.I.C. and T.G. performed analysis of the LCOGT photometric follow-up of the system. A.E. and C.M. applied and analysed the Bern Population Synthesis Models. C.X.H. and L.S. developed and ran the MIT Quick Look Pipeline, which identified the candidate in the TESS data. G.W.K. performed the photoevaporation analysis. J.L.-B. obtained and analysed the Astralux data and synthesized all high-resolution imaging results. E.M. obtained the NaCo imaging data. H.O. contributed to the NCORES HARPS programme and the NGTS survey and helped to create the main figures. J.O., O.M., M.D., R.H., M. Lozovsky and C.D. performed the interior-structure calculations. D.V. performed analysis on the potential for tidal self-disruption. C.Z. obtained the SOAR data and provided text summarising the SOAR results. T.-G.T. obtained a further transit with the PEST telescope. J.J.L. contributed to the internal structure discussion. K.G.S. provided the independent check of stellar parameters. M.B. and S. Gandhi calculated estimates of the required telescope time for atmospheric characterization. D.R.A., M.M., E.M.B., C.A.W., J.S.J., J.I.V., J.S.A., D. Bayliss, C. Belardi, M.R.B., S.L.C., A.C., P.E., S. Gill, M.R.G., M.N.G., M. Lendl, J.M., D.P., D.Q., L.R., R.H.T. and R.G.W. contributed to the NGTS facility, in planning, management, data collection or detrending. D.J.A.B., S.H., D. Barrado, S.C.C.B., P.A.W., L.D.N., D. Bayliss, F.B., B.F.C., R.F.D., O.D.S.D., X.D., P.F., J.J., G.M.K., A.S., S.U., P.A.W., J.M.A. and A.O. contributed to the HARPS large programme under which the HARPS data were obtained. D.R.C., I.J.M.C., J.E.S. and S.B.H. contributed to the NaCo imaging data. C. Briceño, N.L. and A.W.M. contributed to the SOAR imaging data. K.D.C., M.F., J.S.J., E.L.N.J., G.R.R., P.R., S.S., E.T., R.V., J.N.W., J.N.V. and Z.Z. provided essential contributions to the TESS mission, which discovered the candidate. All authors read the manuscript and provided general comments.

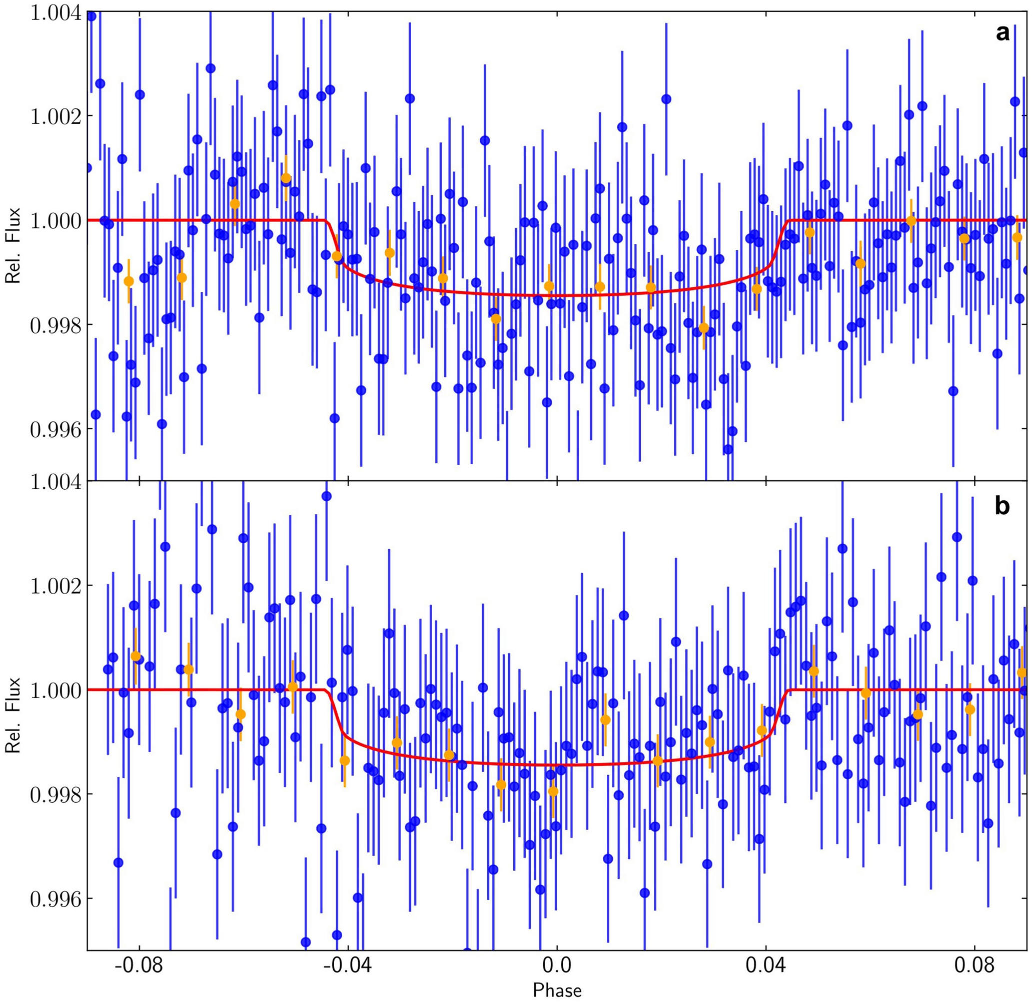
Competing interests The authors declare no competing interests.

Additional information

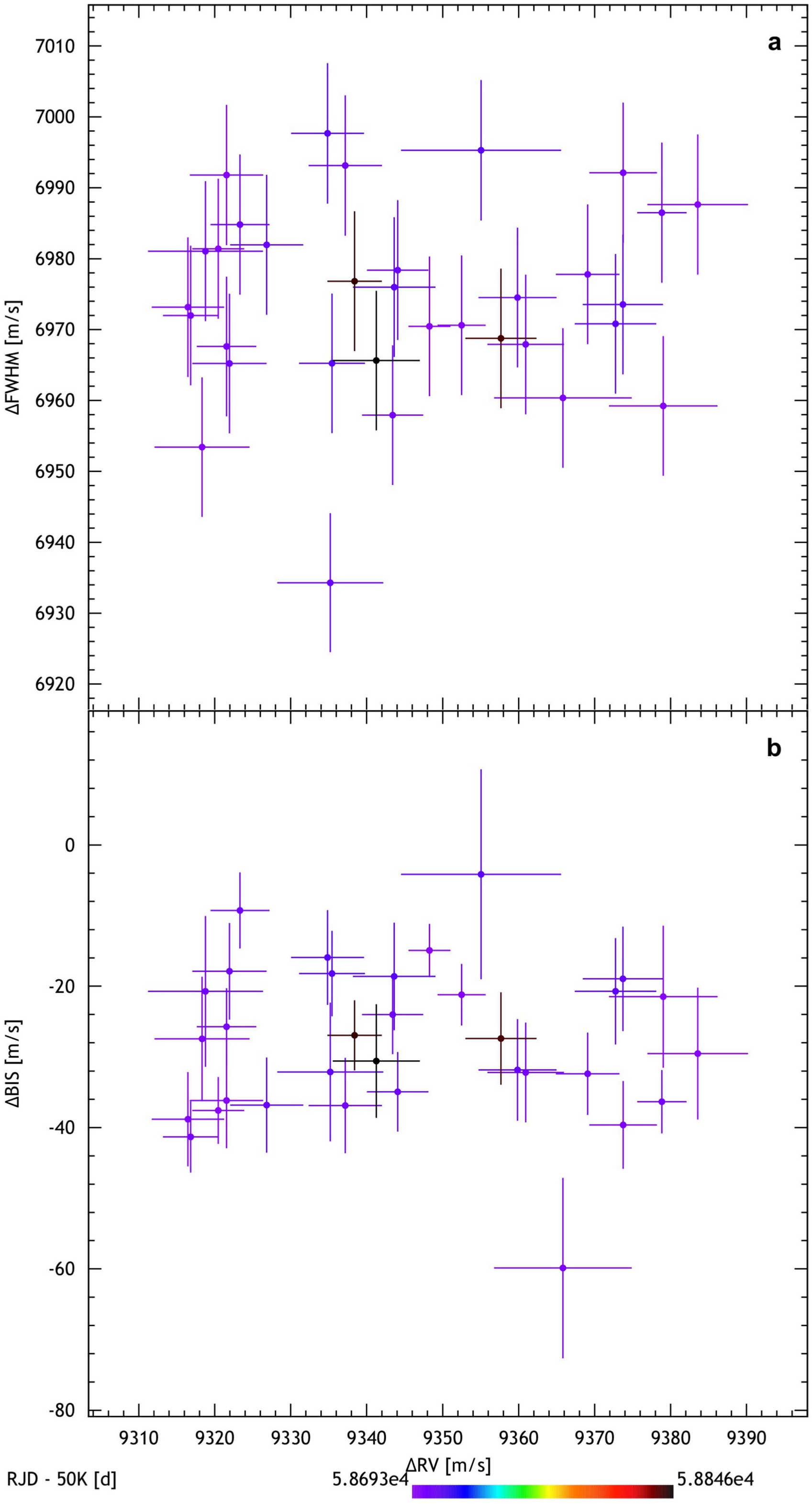
Correspondence and requests for materials should be addressed to D.J.A.

Peer review information *Nature* thanks Roman Baluev and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

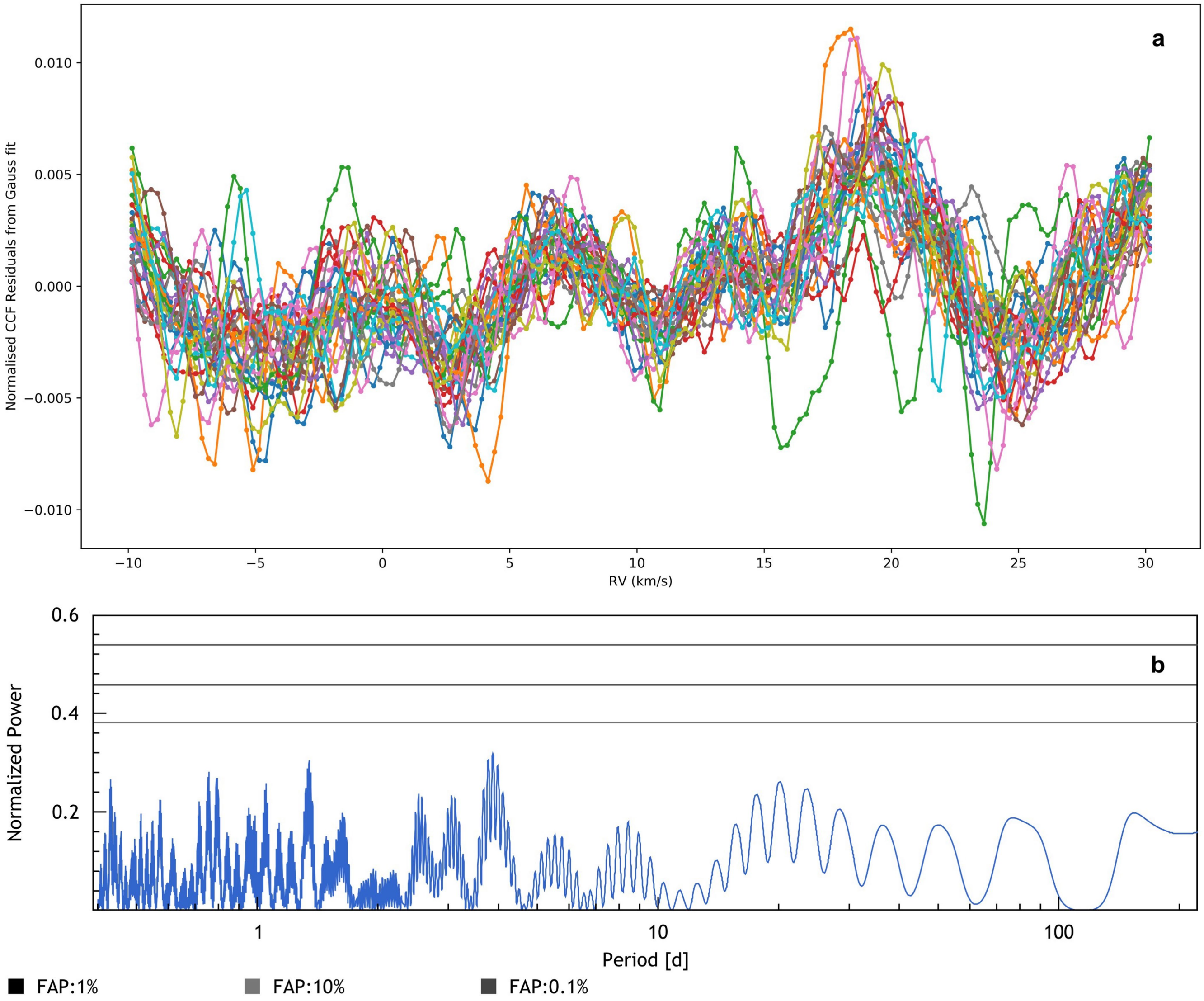
Reprints and permissions information is available at http://www.nature.com/reprints.



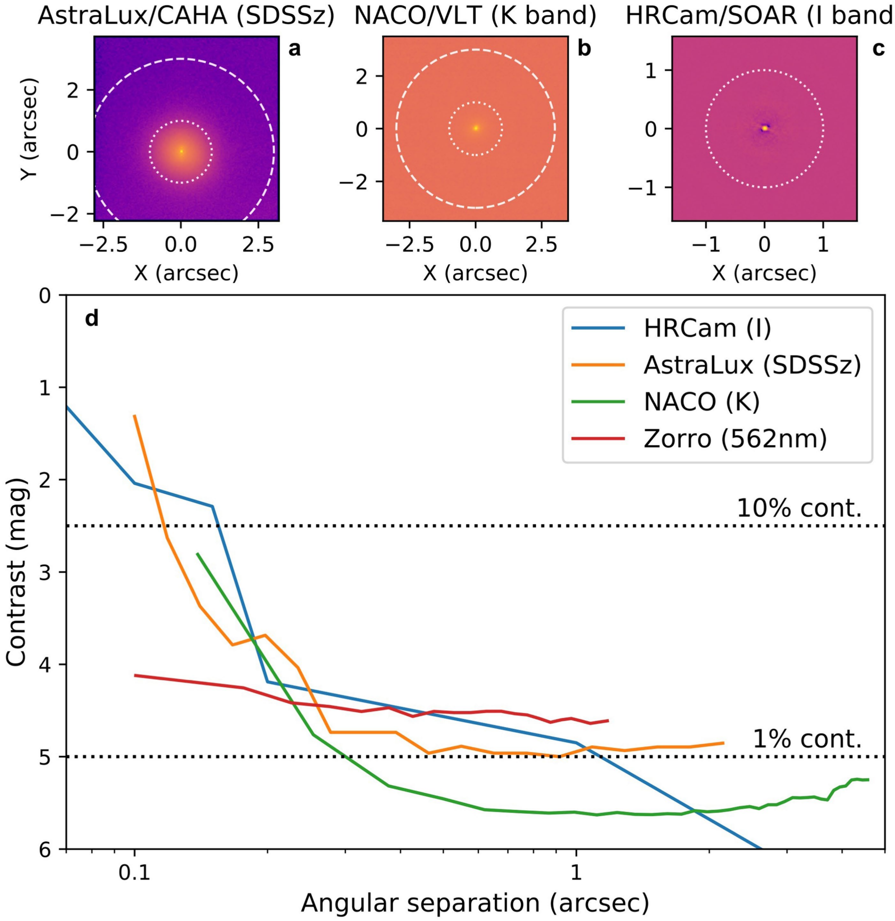
Extended Data Fig. 1 | Photometric data captured by the LCOGT network. **a**, **b**, Data captured on the nights of 30 July 2019 UT (**a**) and 9 August 2019 UT (**b**). The best-fitting model is plotted in red and the binned data in orange. Error bars of individual points show one standard deviation. In the case of binned measurements, points and error bars show the weighted mean and its standard error, respectively.



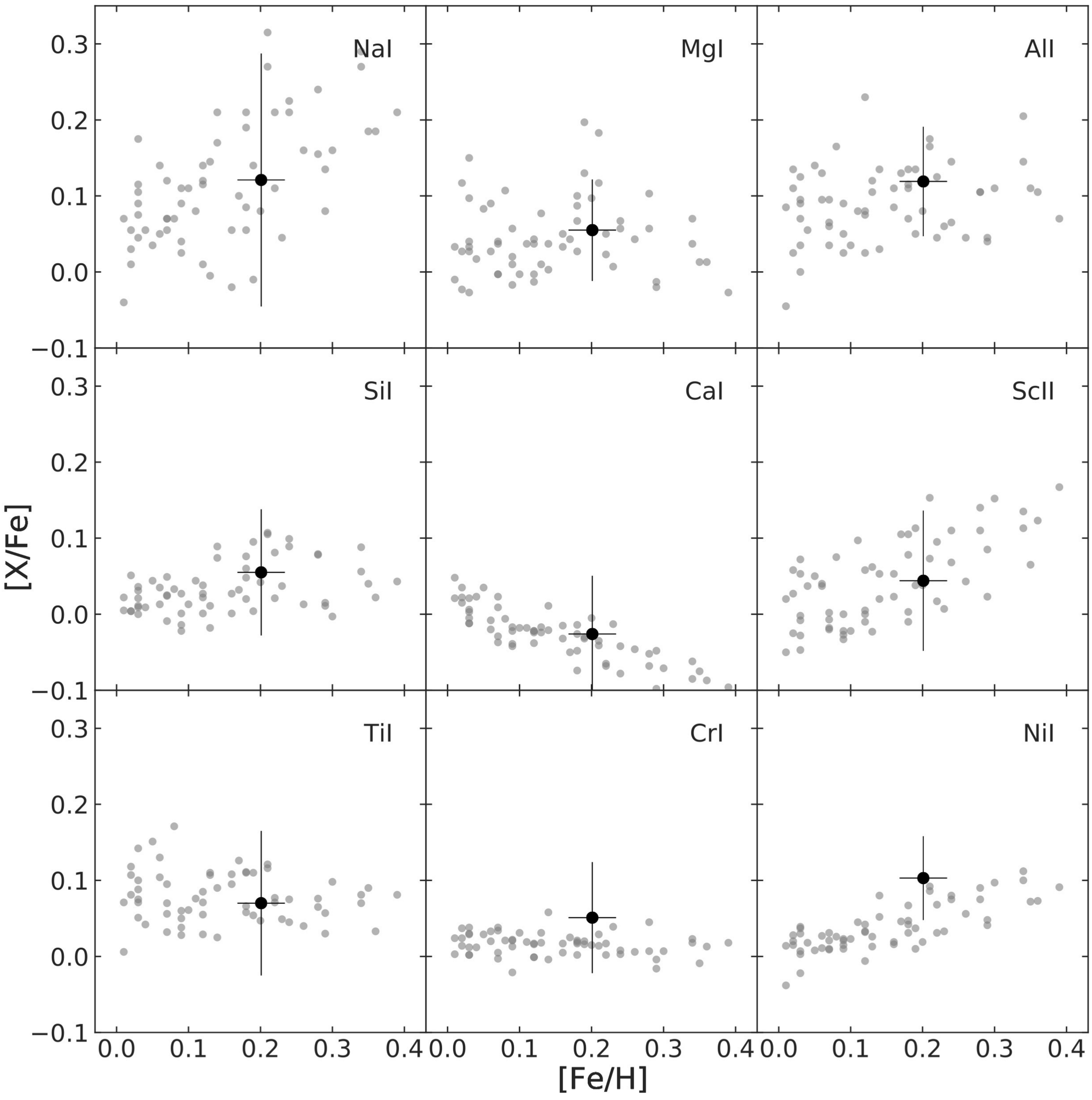
Extended Data Fig. 2 | HARPS activity correlation indicators. **a**, HARPS radial velocities plotted against their bisector value. Colours represent the time of observation measured in BJD−2,400,000. **b**, As for **a**, for the FWHM of the CCF. No correlation is seen in either case. All error bars show one standard deviation.



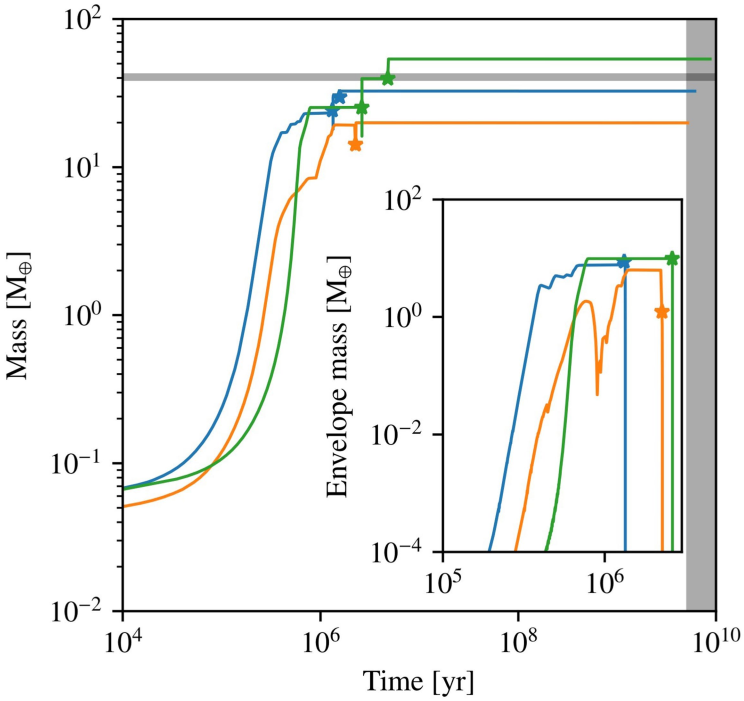
Extended Data Fig. 3 | Tests on the HARPS residuals. a, CCF for the HARPS spectra computed using a G2V template. A Gaussian fit has been removed to leave the residual noise. No clear evidence of a contaminating star is seen. **b**, Periodogram of the HARPS RV residuals. No evidence of periodic structure is found. FAP represents false-alarm probability.



Extended Data Fig. 4 | Collected high-resolution imaging results from AstraLux/CAHA, NaCo/VLT, HRCam/SOAR and Zorro/Gemini (562 nm). **a**–**c**, Images from AstraLux (**a**), NaCo (**b**) and HRCam (**c**). **d**, Sensitivity curves for **a**–**c** and the Zorro 562-nm observation. Our simultaneous 832-nm Zorro observation provides a similar result. The 1% and 10% contrast curves are also plotted.



Extended Data Fig. 5 | TOI-849 compared to field stars. Abundance ratio [X/Fe] against stellar metallicity for TOI-849 (black) and for field stars from the HARPS sample (grey) with similar stellar parameters: $T_{\mathrm{eff}} = 5,329 \pm 200$ K, $\log g = 4.28 \pm 0.20$ dex and [Fe/H] = +0.20 ± 0.20 dex. All error bars show one standard deviation.



Extended Data Fig. 6 | Planet mass against time for three similar planets to TOI-849b in the Bern Population Synthesis models. Grey shaded regions mark the parameters of TOI-849b. Stars mark the time of a giant impact. The inset shows the envelope mass, which is removed after collision.

Extended Data Table 1 | List of stellar and planetary parameters used in the analysis

Parameter	Prior	Posterior	
		Dartmouth (adopted)	PARSEC
Stellar Parameters			
Effective temperature $T_{\rm eff}$ [K]	$\mathcal{N}(5329.0, 48.0)$	$5373.8^{+42.5}_{-41.2}$	$5377.0^{+40.5}_{-39.9}$
Surface gravity log g [cgs]	$\mathcal{N}(4.43, 0.3)$	$4.48^{+0.03}_{-0.04}$	$4.47^{+0.03}_{-0.04}$
Iron abundance [Fe/H] [dex]	$\mathcal{N}(0.201, 0.033)$	0.19 ± 0.03	0.2 ± 0.03
Distance to Earth D [pc]	$\mathcal{N}(224.56, 7.1)$	$224.9^{+5.9}_{-5.5}$	$224.7^{+6.2}_{-5.8}$
Interstellar extinction $E(B-V)$ [mag]	$\mathcal{U}(0.0, 1.0)$	$0.011^{+0.016}_{-0.008}$	$0.01^{+0.016}_{-0.008}$
Systemic radial velocity γ [km s^{-1}]	$\mathcal{U}(5.0, 15.0)$	$9.3502^{+0.0014}_{-0.0013}$	$9.3503^{+0.0014}_{-0.0013}$
Limb-darkening u_a	(derived)	$0.3766^{+0.007}_{-0.0071}$	$0.376^{+0.0069}_{-0.0071}$
Limb-darkening u_b	(derived)	$0.2385^{+0.0041}_{-0.004}$	$0.2389^{+0.004}_{-0.0039}$
Stellar density $\rho_\star/\rho_\odot$	(derived)	$1.195^{+0.118}_{-0.123}$	$1.175^{+0.118}_{-0.140}$
Stellar mass $M_\star$ [$M_\odot$]	(derived)	$0.929^{+0.023}_{-0.023}$	$0.901^{+0.027}_{-0.026}$
Stellar radius $R_\star$ [$R_\odot$]	(derived)	$0.919^{+0.029}_{-0.023}$	$0.916^{+0.031}_{-0.023}$
Stellar age τ [Gyr]	(derived)	$6.7^{+2.9}_{-2.4}$	$8.6^{+3.4}_{-2.9}$
Planet b Parameters			
Orbital Period P_b [d]	$\mathcal{N}(0.76552484, 4.35e-06)$	$0.76552414^{+0.00000262}_{-0.00000279}$	$0.76552402^{+0.00000262}_{-0.00000280}$
Epoch $T_{0,b}$ [BJD - 2450000]	$\mathcal{N}(8394.73741796, 0.0017159129)$	$8394.73768^{+0.00091}_{-0.00099}$	$8394.73768^{+0.00098}_{-0.00096}$
RV semi-amplitude K_b [km s^{-1}]	$\mathcal{U}(0.0, 0.1)$	$0.02864^{+0.00187}_{-0.00182}$	$0.02862^{+0.00198}_{-0.00179}$
Orbital inclination i_b [°]	$\mathcal{S}(50.0, 90.0)$	$86.8^{+2.1}_{-2.4}$	$86.4^{+2.4}_{-2.7}$
Planet-to-star radius ratio k_b	$\mathcal{U}(0.0, 1.0)$	$0.03443^{+0.00092}_{-0.00088}$	$0.03445^{+0.00093}_{-0.00093}$
Orbital eccentricity e_b	$\mathcal{T}(0.0, 0.083, 0.0, 1.0)$	0.0 ± 0.0	0.0 ± 0.0
Argument of periastron ω_b [°]	$\mathcal{U}(0.0, 360.0)$	0.0 ± 0.0	0.0 ± 0.0
System scale $a_b/R_\star$	(derived)	$3.7^{+0.1}_{-0.1}$	$3.7^{+0.1}_{-0.2}$
Impact parameter b_b	(derived)	$0.212^{+0.145}_{-0.140}$	$0.233^{+0.158}_{-0.156}$
Transit duration $T_{14,b}$ [h]	(derived)	1.57 ± 0.04	1.57 ± 0.04
Semi-major axis a_b [AU]	(derived)	$0.01598^{+0.00013}_{-0.00014}$	$0.01582^{+0.00016}_{-0.00015}$
Planet mass M_b [$M_\oplus$]	(derived)	$39.09^{+2.66}_{-2.55}$	$38.33^{+2.70}_{-2.53}$
Planet radius R_b [$R_\oplus$]	(derived)	$3.444^{+0.157}_{-0.115}$	$3.435^{+0.179}_{-0.130}$
Planet bulk density ρ_b [g cm^{-3}]	(derived)	$5.2^{+0.7}_{-0.8}$	$5.2^{+0.8}_{-0.8}$

The respective priors are provided together with the posteriors for the Dartmouth and PARSEC stellar evolution tracks. The posterior values represent the median and 68.3% credible interval. Derived values that might be useful for follow-up work are also reported. $\mathcal{N}(\mu, \sigma^2)$, normal distribution with mean μ and width σ^2; $\mathcal{U}(a, b)$, uniform distribution between a and b; $\mathcal{S}(a, b)$, sine distribution between a and b ; $\mathcal{T}(\mu, \sigma^2, a, b)$, truncated normal distribution between a and b with mean μ and width σ^2.

Extended Data Table 2 | List of instrument parameters used in the analysis

Parameter	Prior	Posterior	
		Dartmouth (adopted)	PARSEC
Instrument-related Parameters			
HARPS jitter $\sigma_{j,\ \mathrm{RV}}$ [km s^{-1}]	$\mathcal{U}(0.0, 0.1)$	$0.00525^{+0.00137}_{-0.00125}$	$0.00525^{+0.00141}_{-0.00122}$
HARPS drift [km s$^{-1}.d^{-1}$]	$\mathcal{U}(-0.001, 0.001)$	$-0.000043^{\pm}0.000031$	-0.000042 ± 0.000030
TESS contamination [%]	$\mathcal{T}(0.0, 0.005, 0.0, 1.0)$	$0.004^{+0.003}_{-0.003}$	$0.003^{+0.004}_{-0.002}$
TESS jitter $\sigma_{j,\ \mathit{TESS}}$ [ppm]	$\mathcal{U}(0.0, 10^5)$	$53.0^{+51.8}_{-36.1}$	$53.0^{+52.5}_{-37.8}$
TESS out-of-transit flux	$\mathcal{U}(0.99, 1.01)$	$1.0001003^{+0.0000209}_{-0.0000204}$	$1.0001001^{+0.0000218}_{-0.0000212}$
TESS limb-darkening u_a	(derived)	0.3766 ± 0.0071	0.3760 ± 0.0071
TESS limb-darkening u_b	(derived)	0.2385 ± 0.0041	0.2389 ± 0.0040
$NGTS_1$ contamination [%]	$\mathcal{T}(0.0, 0.005, 0.0, 1.0)$	$0.003^{+0.004}_{-0.002}$	$0.003^{+0.004}_{-0.002}$
$NGTS_1$ jitter $\sigma_{j,\ \mathit{NGTSfirst}}$ [ppm]	$\mathcal{U}(0.0, 10^5)$	$71.9^{+84.2}_{-51.7}$	$74.7^{+82.4}_{-52.7}$
$NGTS_1$ out-of-transit flux	$\mathcal{U}(0.99, 1.01)$	$1.0000853^{+0.0000801}_{-0.0000896}$	$1.00008^{+0.0000864}_{-0.0000862}$
$NGTS_2$ contamination [%]	$\mathcal{T}(0.0, 0.005, 0.0, 1.0)$	$0.003^{+0.003}_{-0.002}$	$0.003^{+0.004}_{-0.002}$
$NGTS_2$ jitter $\sigma_{j,\ \mathit{NGTSsecond}}$ [ppm]	$\mathcal{U}(0.0, 10^5)$	$85.1^{+98.7}_{-59.7}$	$84.8^{+97.2}_{-60.2}$
$NGTS_2$ out-of-transit flux	$\mathcal{U}(0.99, 1.01)$	$1.0000869^{+0.0000839}_{-0.0000962}$	$1.0000765^{+0.0000963}_{-0.0000944}$
NGTS limb-darkening u_a	(derived)	0.4758 ± 0.0080	0.4752 ± 0.0081
NGTS limb-darkening u_b	(derived)	0.2114 ± 0.0050	0.2118 ± 0.0050
LCO_1 contamination [%]	$\mathcal{T}(0.0, 0.005, 0.0, 1.0)$	$0.003^{+0.004}_{-0.002}$	$0.003^{+0.004}_{-0.002}$
LCO_1 jitter $\sigma_{j,\ \mathit{LCOsecond}}$ [ppm]	$\mathcal{U}(0.0, 10^5)$	$1022.0^{+85.6}_{-86.9}$	$1022.6^{+88.3}_{-90.2}$
LCO_1 out-of-transit flux	$\mathcal{U}(0.98, 1.02)$	$0.9999986^{+0.0000880}_{-0.0000870}$	$0.9999953^{+0.0000870}_{-0.0000879}$
LCO_2 contamination [%]	$\mathcal{T}(0.0, 0.005, 0.0, 1.0)$	$0.003^{+0.004}_{-0.002}$	$0.003^{+0.004}_{-0.002}$
LCO_2 jitter $\sigma_{j,\ \mathit{LCOthird}}$ [ppm]	$\mathcal{U}(0.0, 10^5)$	$1417.1^{+86.7}_{-76.3}$	$1420.7^{+86.7}_{-84.7}$
LCO_2 out-of-transit flux	$\mathcal{U}(0.98, 1.02)$	$0.9999873^{+0.0000947}_{-0.0000976}$	$0.9999938^{+0.0001048}_{-0.0001031}$
LCO limb-darkening u_a	(derived)	0.3828 ± 0.0072	0.3822 ± 0.0073
LCO limb-darkening u_b	(derived)	0.2386 ± 0.0044	0.2390 ± 0.0042
SED jitter [mag]	$\mathcal{U}(0.0, 0.1)$	$0.049^{+0.03}_{-0.026}$	$0.047^{+0.031}_{-0.026}$

The respective priors are provided together with the posteriors for the Dartmouth and PARSEC stellar evolution tracks. The posterior values represent the median and 68.3% credible interval. $\mathcal{N}(\mu, \sigma^2)$, normal distribution with mean μ and width σ^2; $\mathcal{U}(a, b)$, uniform distribution between *a* and *b*; $\mathcal{T}(\mu, \sigma^2, a, b)$, truncated normal distribution between *a* and *b* with mean μ and width σ^2.

Extended Data Table 3 | Stellar properties of TOI-849

Property	Value	Source
Astrometric Properties		
RA	01:54:51.7910	GAIA DR2
Dec	-29:25:18.1508	GAIA DR2
TIC ID	33595516	TICv8
GAIA ID	5023809953208388352	GAIA DR2
2MASS ID	01545169-2925186	2MASS
μ_{RA} (mas.yr^{-1})	73.315	GAIA DR2
μ_{Dec} (mas.yr^{-1})	20.664	GAIA DR2
Photometric Properties		
TESS (mag)	11.55	TICv8
B (mag)	12.84	TICv8
V (mag)	11.98	TICv8
G (mag)	12.06	TICv8
J (mag)	10.83	TICv8
H (mag)	10.48	TICv8
K (mag)	10.42	TICv8

Sources: Gaia Data Release 2[10], TICv8[139], 2MASS[140].

Extended Data Table 4 | HARPS radial velocities

BJD d	RV kms^{-1}	σ_{RV} kms^{-1}	CCF FWHM kms^{-1}	CCF Contrast	Bisector kms^{-1}	S/N(50)	$\mathrm{T}_{\mathrm{exp}}$ s	$\mathrm{Airmass}_{\mathrm{start}}$
2458692.78910182	9.320464	0.003341	6.9814	57.568	-0.0376	28.9	1800	1.464
2458692.87197814	9.348264	0.002660	6.9705	57.606	-0.0149	35.0	1800	1.073
2458693.78193905	9.379034	0.007109	6.9592	58.265	-0.0215	16.7	1500	1.49
2458693.86713031	9.383575	0.006590	6.9876	57.935	-0.0295	17.6	1200	1.069
2458694.79485824	9.352493	0.003080	6.9706	57.677	-0.0212	31.2	1500	1.365
2458694.89266609	9.321555	0.004772	6.9918	57.768	-0.0362	22.3	1200	1.022
2458695.76161626	9.316472	0.004716	6.9732	57.804	-0.0388	22.9	1500	1.633
2458695.8578846	9.318335	0.006212	6.9534	58.095	-0.0274	18.1	1200	1.078
2458697.77111987	9.373754	0.004389	6.9921	57.730	-0.0396	24.0	1800	1.508
2458697.86515415	9.365826	0.009028	6.9604	58.639	-0.0599	13.5	1200	1.051
2458698.79553074	9.316836	0.003574	6.9720	57.644	-0.0413	28.1	1500	1.296
2458698.86273442	9.321557	0.003851	6.9676	57.785	-0.0257	26.1	1500	1.055
2458699.77215996	9.343418	0.003958	6.9579	57.855	-0.0240	26.0	1500	1.434
2458699.86782619	9.360927	0.004984	6.9679	57.876	-0.0322	21.5	1200	1.038
2458700.7860712	9.378852	0.003165	6.9865	57.620	-0.0363	30.7	1500	1.321
2458700.86459501	9.369082	0.004117	6.9778	57.755	-0.0324	24.9	1200	1.039
2458701.74930712	9.337185	0.004769	6.9931	57.629	-0.0369	22.8	1200	1.573
2458701.82063133	9.321939	0.004835	6.9652	57.948	-0.0179	21.9	1200	1.139
2458701.91235041	9.318780	0.007543	6.9811	58.274	-0.0207	15.3	1200	1.0
2458702.75424659	9.323328	0.003815	6.9848	57.651	-0.0093	26.7	1200	1.504
2458702.82285066	9.344083	0.003979	6.9784	57.701	-0.0349	25.8	1200	1.125
2458705.75330754	9.335210	0.006936	6.9343	57.874	-0.0321	16.5	1200	1.443
2458705.8276873	9.326840	0.004757	6.9820	57.743	-0.0368	22.3	1200	1.087
2458705.92257763	9.359851	0.005083	6.9745	57.812	-0.0318	21.9	1200	1.007
2458706.89905173	9.373723	0.005230	6.9735	57.873	-0.0190	20.6	1800	1.0
2458707.74440581	9.372752	0.005320	6.9708	57.367	-0.0207	20.7	1800	1.508
2458707.85009529	9.355053	0.010510	6.9953	57.194	-0.0042	12.4	1800	1.036
2458708.72594834	9.334852	0.004742	6.9977	57.107	-0.0159	22.9	1200	1.626
2458708.82117413	9.335439	0.004273	6.9652	57.287	-0.0182	24.4	1200	1.084
2458708.92270817	9.343619	0.005383	6.9760	57.703	-0.0186	20.9	1200	1.013
2458838.6181445	9.357671	0.004621	6.9688	57.643	-0.0274	22.8	1800	1.096
2458840.60848862	9.338411	0.003501	6.9768	57.653	-0.0269	28.3	1800	1.087
2458845.6514276	9.341266	0.005679	6.9656	57.812	-0.0306	20.6	1800	1.322

S/N represents the signal-to-noise ratio. σ_{RV} is the 1σ error on the RV measurements. $\mathrm{T}_{\mathrm{exp}}$ is the exposure time of the observation.

Quantum correlations between light and the kilogram-mass mirrors of LIGO

https://doi.org/10.1038/s41586-020-2420-8

Received: 3 February 2020

Accepted: 4 May 2020

Haocun Yu[1]✉, L. McCuller[1]✉, M. Tse[1], N. Kijbunchoo[2], L. Barsotti[1], N. Mavalvala[1] and other members of the LIGO Scientific Collaboration*

The measurement of minuscule forces and displacements with ever greater precision is inhibited by the Heisenberg uncertainty principle, which imposes a limit to the precision with which the position of an object can be measured continuously, known as the standard quantum limit[1–4]. When light is used as the probe, the standard quantum limit arises from the balance between the uncertainties of the photon radiation pressure applied to the object and of the photon number in the photoelectric detection. The only way to surpass the standard quantum limit is by introducing correlations between the position/momentum uncertainty of the object and the photon number/phase uncertainty of the light that it reflects[5]. Here we confirm experimentally the theoretical prediction[5] that this type of quantum correlation is naturally produced in the Laser Interferometer Gravitational-wave Observatory (LIGO). We characterize and compare noise spectra taken without squeezing and with squeezed vacuum states injected at varying quadrature angles. After subtracting classical noise, our measurements show that the quantum mechanical uncertainties in the phases of the 200-kilowatt laser beams and in the positions of the 40-kilogram mirrors of the Advanced LIGO detectors yield a joint quantum uncertainty that is a factor of 1.4 (3 decibels) below the standard quantum limit. We anticipate that the use of quantum correlations will improve not only the observation of gravitational waves, but also more broadly future quantum noise-limited measurements.

The Heisenberg uncertainty principle dictates that once an object is localized with sufficient precision, the momentum of that object must become accordingly uncertain. In a one-off measurement, this does not pose a problem. However, when the position of an object must be measured continuously, as in gravitational wave (GW) detectors, the momentum uncertainty introduced by the act of measuring the position evolves into a position uncertainty for future position measurements—a process known as quantum backaction. In striking a balance between the precision of position measurements and the imprecision caused by quantum backaction, an apparent maximum precision is reached for a continuous position measurement. This is the standard quantum limit (SQL), and for an interferometric measurement, as long as the shot noise and quantum radiation pressure noise (QRPN) are uncorrelated, the SQL is indeed the limit.

The SQL was first introduced by Braginsky et al.[2,3] as a fundamental limit to the sensitivity of GW detectors. It should be possible to reach the SQL with objects that are macroscopic or even human-scale because the quantization of the probe light is what enforces the SQL (see, for example, footnote 1 of ref. [4]). In principle, the SQL can be surpassed when the shot noise and the QRPN are correlated. Such correlations already exist in the interferometer because incoming quantum fluctuations entering from its output port drive both the shot noise and the QRPN, giving rise to ponderomotive squeezing. An injected squeezed state, when combined appropriately with ponderomotive squeezing, enables surpassing the SQL (see section IVB of ref. [4]). Alternative methods for surpassing the SQL in GW detectors are presented in refs. [4,6].

Here, we inject a laser mode that is in a squeezed vacuum state into a laser interferometric GW detector with 40-kg mirrors, and use the optomechanically induced correlations of ponderomotive squeezing to demonstrate quantum noise below the SQL. This measurement marks two milestones of quantum measurement. First, we directly observe the contribution of the QRPN to the motion of kilogram-mass objects at room temperature, indicating that quantum backaction imposed by the Heisenberg uncertainty principle persists even at human scales. Second, we demonstrate quantum noise below the SQL, proving the existence of quantum correlations involving the position uncertainty of the 40-kg mirrors. This measurement is an important step towards further improvements in GW sensitivity through quantum engineering techniques[4,6–10].

A considerable barrier to revealing quantum correlations between light and macroscopic objects is the ubiquitous presence of thermal fluctuations that drive their motion. Previous demonstrations of the QRPN have involved cryogenically precooled pico- to microgram-scale mechanics[10–14], with three exceptions[15–17]. Similarly, a previous sub-SQL measurement of displacement was performed on a cryogenically precooled mechanical oscillator at the nanogram

---

[1]LIGO, Massachusetts Institute of Technology, Cambridge, MA, USA. [2]OzGrav, Australian National University, Canberra, Australian Capital Territory, Australia. *A list of authors and their affiliations appears at the end of the paper. ✉e-mail: haocunyu@mit.edu; mcculler@mit.edu

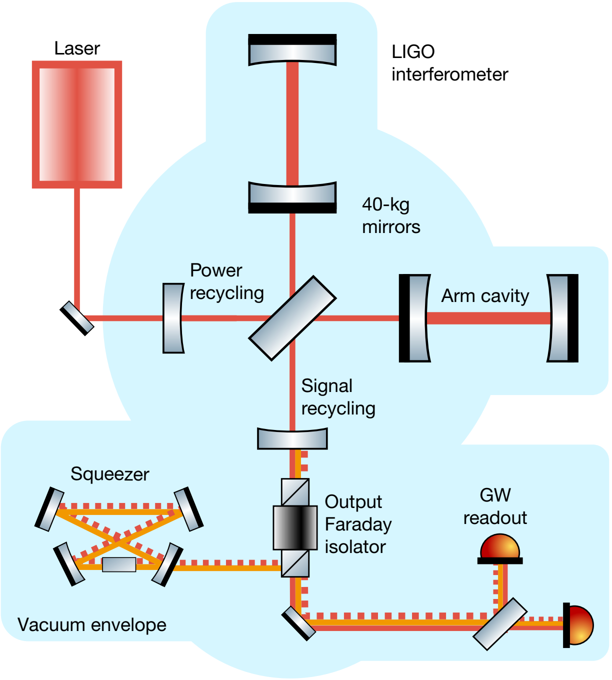


Fig. 1 | Simplified schematic of the experimental setup. Squeezed vacuum (dashed red line) is injected through the output Faraday isolator and co-propagates with the 1,064-nm light (solid red line) of the main interferometer. A frequency-shifted control field (orange line) is used to sense the squeeze angle and control it using the phase of the squeezer pump field (not shown)[19].

mass scale[18]. The measurements presented here are performed on the room-temperature, 40-kg mirrors of Advanced LIGO using laser light of 200 kW, and are enabled by the injection of squeezed states and sufficiently low classical noise. The classical noise is subtracted to reveal quantum noise below the SQL.

We performed this experiment using the Advanced LIGO detector in Livingston, Louisiana, USA. For the third astrophysics observing run of LIGO/Virgo, squeezed vacuum is injected into the interferometer, with the squeezing level and squeezing quadrature angle tuned to maximize the GW sensitivity[19]. In this experiment, the interferometer is maintained in the observing configuration, but data are taken with an increased squeezing level and over a range of squeezing angles in order to fully characterize the quantum noise.

The Advanced LIGO detector is a Michelson interferometer with two 4-km Fabry–Pérot arms, as well as power- and signal-recycling cavities at the input and output ports of the beam splitter, respectively (see Fig. 1). The arm-cavity optics are 40-kg fused-silica mirrors, suspended as pendulums inside an ultrahigh-vacuum envelope[20]. During the measurement, 200 ± 10 kW of 1,064-nm laser power circulates in each arm cavity. The differential arm displacement signal (Δx) is detected as modulations of a small static field at the GW readout caused by a deliberate mismatch in the interferometer arm lengths[20]. The displacement signal Δx is part of a closed servo loop, which is monitored by a continuous calibration procedure that also extracts the instrument sensing function by driving the differential arm motion and measuring the optical response. Details of the squeezed light source and its operation, including the control method for adjusting the squeezing angle, are provided in ref. [19]. For this measurement, injected squeezing results in 3.3 dB of squeezing and 7.7 dB of antisqueezing measured at the GW readout.

An analytic model of the displacement sensitivity in an ideal LIGO interferometer illustrates how the combination of ponderomotive squeezing and injected squeezing allows us to surpass the SQL for the differential arm motion. A model that builds on methods developed in refs. [4,6], with extensions to account for losses and off-resonance cavities, is provided in Methods. Here, the ideal model is used for clarity. The application of the Heisenberg uncertainty principle to an interferometric measurement of differential displacement Δx sets a limit to the one-sided spectral density of:

$$\Delta x^2(\Omega) = S(\Omega,\phi)[1+\mathcal{K}^2(\Omega)]\frac{\hbar c}{8k\,|G(\Omega)|^2 P_{\mathrm{arm}}} \tag{1}$$

with

$$\mathcal{K}(\Omega) = \frac{32k\,|G(\Omega)|^2 P_{\mathrm{arm}}}{m\Omega^2 c},\quad G(\Omega) \equiv \sqrt{\frac{\gamma c}{2L}\,\frac{1}{\gamma + \mathrm{i}\Omega}}. \tag{2}$$

Here P_{arm} is the circulating arm power; k is the laser wavenumber; $\Omega/(2\pi)$ is the sideband frequency of the GW readout; m is the mass of each mirror; $\hbar$ is the reduced Planck constant; L is the arm length of 3,995 m; and γ is the signal bandwidth of $2\pi \times 450$ Hz in LIGO. $G(\Omega)$ is the optical-field transmissivity between the arm cavities and the readout detector, making $2kG(\Omega)\sqrt{P_{\mathrm{arm}}}$ the sensing function that relates Δx to the emitted optical field that modulates the GW readout power.

The factors $S(\Omega,\phi)$ and $1+\mathcal{K}^2(\Omega)$ (ϕ, squeeze angle) capture the radiation pressure interaction, whereby the mirror oscillator motion correlates the injected optical amplitude quadrature to the output phase quadrature, with $\mathcal{K}(\Omega)$ the ponderomotive interaction strength. The theory of ponderomotive squeezing is detailed in section IVA-B of ref. [4]. $S(\Omega,\phi)$ accounts for the injection of squeezed states. Without injected squeezing, $S = 1$, in which case the arm power P_{arm} may be chosen to minimize $\Delta x(\Omega)$ by balancing the shot noise and the radiation pressure noise. The resulting minimum $\Delta x_{\mathrm{SQL}}(\Omega)$ is the free-mass SQL for a Michelson interferometer with a Fabry–Pérot cavity in each arm[4]:

$$\Delta x^2(\Omega) \geq \Delta x^2_{\mathrm{SQL}}(\Omega) \equiv \frac{8\hbar}{m\Omega^2}. \tag{3}$$

When injecting squeezed states at a squeeze angle ϕ with a squeeze factor r, the squeezing measured at the readout, $S(\Omega,\phi)$, becomes:

$$S(\Omega,\phi) = \mathrm{e}^{-2r}\cos^2[\phi-\theta(\Omega)] + \mathrm{e}^{2r}\sin^2[\phi-\theta(\Omega)] \tag{4}$$

$$\theta(\Omega) = \arctan[\mathcal{K}(\Omega)]. \tag{5}$$

$\phi = 0$ is defined as the squeezing angle that reduces the power spectral density of the shot noise, where $\theta \rightarrow 0$, by a factor of e^{-2r}.

The expression $\phi - \theta(\Omega)$ characterizes the frequency-dependent interaction between ponderomotive and injected squeezing. Equation (4) indicates that at frequencies for which $\theta(\Omega) = \phi$, these two combine to produce a minimum in the quantum noise spectrum, which appears as a ‘dip’ in the curves of Fig. 2. Whereas the $S = 1$ case led to the SQL in equation (3), injecting squeezed states allows the SQL to be surpassed at measurement frequencies for which $S(\Omega,\phi) < 1$.

Figure 2 shows amplitude spectral densities of differential displacement. Exposing the sub-SQL dip requires reliably estimating and subtracting classical noise around 40 Hz. The data are acquired as three sets of spectral measurements in each of two operating modes: with and without squeezing injection. By alternating operation between the two modes, we establish that the noise is consistent within statistical variations, confirming that it is stationary over the duration of the experiment. To further address the concern that the classical noise between modes of operation may be changing, additional data at a range of squeezing angles are obtained, as shown in Fig. 3.

In Fig. 2, the black trace is the measured total noise at the readout with squeezing disengaged, including both quantum and classical noise

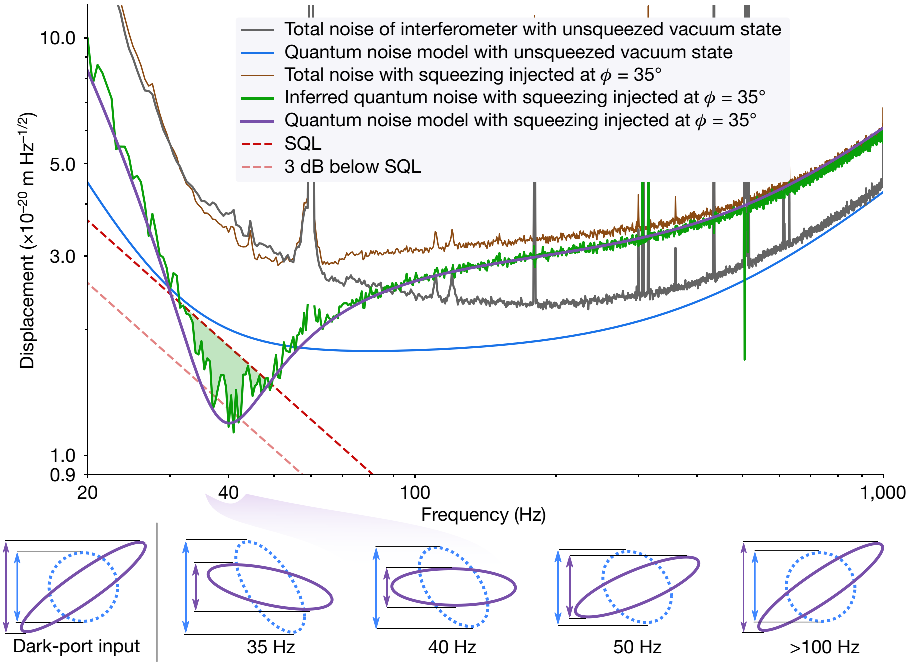


Fig. 2 | Spectral density measurements revealing sub-SQL quantum noise. Top, spectral density of the differential displacement (Δx) noise of the interferometer. The grey and brown traces show the measured total noise level of the interferometer with the unsqueezed vacuum state (that is, the reference) and injected squeezing at 35°, respectively. The blue trace is the model of quantum noise during the reference measurement. The green trace shows the inferred quantum noise of the interferometer with injected squeezing at 35°, and its corresponding model is the purple trace. The notch feature, or 'dip', results from the ponderomotive squeezing affecting the injected optical squeezed states. It reaches −3 dB of the free-mass SQL (red dashed trace; given by equation (3)) at 40 Hz. Bottom, phase-space representation of the modelled quantum states entering through the dark port of the interferometer (left) and the output states (right), with their frequency dependence indicated. Shown are the cases in which the input state is unsqueezed vacuum (dashed blue line) and squeezing at ϕ = 35° (solid purple line). In the unsqueezed vacuum case, ponderomotive squeezing distorts the ellipse for frequencies below 100 Hz, increasing the QRPN in the readout quadrature (blue arrows). In the case of injected squeezing, the same physical process creates a state with reduced noise at 40 Hz (purple arrows).

contributions. It is generated from a 90-min average split across three non-contiguous time periods in which the squeezer cavity is set to be off resonance[19], allowing the unsqueezed vacuum state to enter the interferometer. The blue trace is the modelled quantum noise contribution to the total noise measurement of the black trace. Subtracting the blue trace from the black trace gives the total classical noise contribution. We verify that this classical noise component is stationary and independent of squeezer status (see discussion in the caption of Fig. 3 and details in Methods). The model shows that quantum noise dominates the interferometer sensitivity at high frequencies ($\Omega > \gamma \approx 2\pi \times 450$ Hz), and accounts for 28% of the total measured noise power at 40 Hz. Of the remaining non-quantum noise, 24% is estimated to be coating and thermo-optic noise, with the rest unidentified (A. Buikema et al., manuscript in preparation).

The green trace in Fig. 2 shows the inferred quantum noise spectrum with squeezing injected at ϕ = 35°. This angle, determined from the model fit, places the dip in the frequency region in which the ratio between the total measured reference noise and the SQL curve is minimized. The green trace is calculated as the total measured displacement spectrum while the squeezer is engaged (brown trace), minus the classical noise contribution determined from the reference measurement. The purple trace shows the quantum noise model corresponding to ϕ = 35° squeezing, featuring a dip in the quantum noise that reaches down to 70% or 3 dB of the SQL at 40 Hz.

Squeezing measurements at three additional ϕ values are presented in Fig. 3. They show that the QRPN contributes to the motion of the Advanced LIGO mirrors. At each ϕ, the quantum noise trace is calculated by subtracting the same classical noise contribution (determined

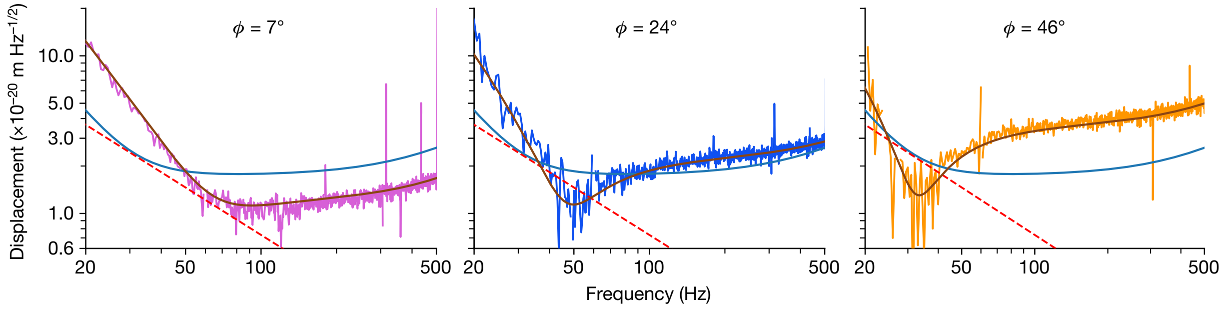


Fig. 3 | Quantum noise spectra at additional squeezing angles of 7°, 24° and 46°. Each dataset is plotted with the same classical noise subtraction as Fig. 2, and with a corresponding quantum noise model curve (copper line). The model without injected squeezing (blue line) is plotted for comparison. The differences between the squeezed datasets and the reference model show that the QRPN contributes to the motion of the Advanced LIGO mirrors. The QRPN contribution can be increased and decreased as the injected state is varied. These data were obtained with less observing time than Fig. 2 and have correspondingly larger statistical fluctuations.

from the reference data) from the measured displacement spectrum. We note that the modelled quantum noise plotted here requires the full functional form of $S(\Omega, \phi, \psi)$ in equation (9) in Methods, rather than the simplified version of equation (4). A total of 12 squeezing measurements are combined to plot $S(\Omega, \phi, \psi)$ in Extended Data Fig. 2.

The uncertainties in both the data and the model are discussed here, with additional details in Methods. The statistical error in the power spectrum measurement of the quantum noise, after subtraction, is 8% at 40 Hz (for a bin width of 0.5 Hz). We test for discrepancies between the three reference datasets and find that the relative uncertainty in the classical noise stationarity is bounded by the same statistical error. Errors in the optical sensing function $2kG(\Omega)\sqrt{P_{\mathrm{arm}}}$, along with the Δx servo loop compensation, are determined from the online interferometer calibration procedure[21] to be ±3%. The uncertainty in the arm-cavity power is 5%. Aside from the reference datasets, the model curves of Figs. 2, 3 require the squeeze factor r and the interferometer losses[19], which are determined from fits across all datasets. The 12 measurements also constrain an additional unwanted frequency-dependent squeezing phase shift of $\psi = 8°$, which accumulates across the frequency region where $\Omega \approx \gamma$. This effect arises from a detuning of the signal-recycling cavity, which is detailed in the Methods, equation (10).

The measurements presented here represent long-awaited milestones in verifying the role of quantum mechanics in limiting the precision of position measurements even for macroscopic objects, and thereby limiting the sensitivity of GW detectors.

First, we observe that the QRPN contributes to the motion of the kilogram-scale mirrors of LIGO. This observation is also made with the Advanced Virgo GW detector (F. Acernese et al., manuscript in preparation). It is remarkable that quantum vacuum fluctuations can influence the motion of these macroscopic, human-scale objects, and that the effect is measured—this is experimental quantum mechanics at its most macroscopic scale.

Second, revealing quantum noise below the SQL in the Advanced LIGO detector is the first realization of a quantum non-demolition technique in GW detectors[2,3], where quantum correlations prevent the measurement device from demolishing the same information that one is trying to extract. Exploiting quantum correlations allows a fundamental quantum limit to be manipulated to improve measurement precision.

Finally, we must not forget the foremost scientific objectives of the Advanced LIGO detectors: they are designed for astrophysical observations of GWs from violent cosmic events. During the third observing run of LIGO/Virgo, the squeezing angle in LIGO is set to optimize the sensitivity of the detectors to GWs from binary neutron star mergers[19]. This is not the squeeze angle at which shot noise is minimized, but that for which the combination of shot noise and QRPN are minimized, implying that backaction evasion plays a role in optimizing the sensitivity of the Advanced LIGO detector. This is one of the factors that has allowed Advanced LIGO to go from detecting roughly one astrophysical event per month in observing runs 1 and 2, to about one astrophysical trigger per week in the third observing run of LIGO/Virgo. With further mitigation of classical noise, the sub-SQL performance of GW detectors promises ever greater astrophysical reach in the future.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2420-8.

1. Caves, C. M. Quantum-mechanical noise in an interferometer. *Phys. Rev. D* **23**, 1693–1708 (1981).
2. Braginsky, V. B. & Khalili, F. Y. Quantum nondemolition measurements: the route from toys to tools. *Rev. Mod. Phys.* **68**, 1–11 (1996).
3. Braginsky, V. B., Khalili, F. Y. & Thorne, K. S. *Quantum Measurement* (Cambridge University Press, 1992).
4. Kimble, H. J., Levin, Y., Matsko, A. B., Thorne, K. S. & Vyatchanin, S. P. Conversion of conventional gravitational-wave interferometers into quantum nondemolition interferometers by modifying their input and/or output optics. *Phys. Rev. D* **65**, 022002 (2001).
5. Unruh, W. G. *Quantum Optics, Experimental Gravitation, and Measurement Theory* (Plenum, 1982).
6. Buonanno, A. & Chen, Y. Quantum noise in second generation, signal-recycled laser interferometric gravitational-wave detectors. *Phys. Rev. D* **64**, 042006 (2001).
7. Kwee, P., Miller, J., Isogai, T., Barsotti, L. & Evans, M. Decoherence and degradation of squeezed states in quantum filter cavities. *Phys. Rev. D* **90**, 062006 (2014).
8. Danilishin, S. et al. Creation of a quantum oscillator by classical control. Preprint at https://arxiv.org/abs/0809.2024 (2008).
9. Purdue, P. & Chen, Y. Practical speed meter designs for quantum nondemolition gravitational-wave interferometers. *Phys. Rev. D* **66**, 122004 (2002).
10. Møller, C. B. et al. Quantum back-action-evading measurement of motion in a negative mass reference frame. *Nature* **547**, 191–195 (2017).
11. Purdy, T. P., Peterson, R. W. & Regal, C. A. Observation of radiation pressure shot noise on a macroscopic object. *Science* **339**, 801–804 (2013).
12. Suh, J. et al. Mechanically detecting and avoiding the quantum fluctuations of a microwave field. *Science* **344**, 1262–1265 (2014).
13. Wilson, D. J. et al. Measurement-based control of a mechanical oscillator at its thermal decoherence rate. *Nature* **524**, 325–329 (2015).
14. Teufel, J., Lecocq, F. & Simmonds, R. Overwhelming thermomechanical motion with microwave radiation pressure shot noise. *Phys. Rev. Lett.* **116**, 013602 (2016).
15. Cripe, J. et al. Measurement of quantum back action in the audio band at room temperature. *Nature* **568**, 364–367 (2019).
16. Sudhir, V. et al. Quantum correlations of light from a room-temperature mechanical oscillator. *Phys. Rev. X* **7**, 031055 (2017).
17. Purdy, T. P., Grutter, K. E., Srinivasan, K. & Taylor, J. M. Quantum correlations from a room-temperature optomechanical cavity. *Science* **356**, 1265–1268 (2017).
18. Mason, D., Chen, J., Rossi, M., Tsaturyan, Y. & Schliesser, A. Continuous force and displacement measurement below the standard quantum limit. *Nat. Phys.* **15**, 745–749 (2019).
19. Tse, M. et al. Quantum-enhanced advanced LIGO detectors in the era of gravitational-wave astronomy. *Phys. Rev. Lett.* **123**, 231107 (2019).
20. Abbott, B. P. et al. Gw150914: the advanced LIGO detectors in the era of first discoveries. *Phys. Rev. Lett.* **116**, 131103 (2016).
21. Sun, L. et al. Characterization of systematic error in advanced LIGO calibration. Preprint at https://arxiv.org/abs/2005.02531 (2020).

Detector group of The LIGO Scientific Collaboration

J. Betzwieser[3], C. D. Blair[3], S. E. Dwyer[4], A. Effler[3], M. Evans[1], A. Fernandez-Galiana[1], P. Fritschel[1], V. V. Frolov[3], F. Matichard[1,5], D. E. McClelland[2], T. McRae[2], A. Mullavey[3], D. Sigg[4], B. J. J. Slagmolen[2], C. Whittle[1], A. Buikema[1], Y. Chen[6], T. R. Corbitt[7], R. Schnabel[8], R. Abbott[5], C. Adams[3], R. X. Adhikari[5], A. Ananyeva[5], S. Appert[5], K. Arai[5], J. S. Areeda[9], Y. Asali[10], S. M. Aston[3], C. Austin[7], A. M. Baer[11], M. Ball[12], S. W. Ballmer[13], S. Banagiri[14], D. Barker[4], J. Bartlett[4], B. K. Berger[15], D. Bhattacharjee[16], G. Billingsley[5], S. Biscans[1,5], R. M. Blair[4], N. Bode[17,18], P. Booker[17,18], R. Bork[5], A. Bramley[3], A. F. Brooks[5], D. D. Brown[19], C. Cahillane[5], K. C. Cannon[20], X. Chen[21], A. A. Ciobanu[19], F. Clara[4], S. J. Cooper[22], K. R. Corley[10], S. T. Countryman[10], P. B. Covas[23], D. C. Coyne[5], L. E. H. Datrier[24], D. Davis[13], C. Di Fronzo[22], K. L. Dooley[25,26], J. C. Driggers[4], P. Dupej[24], T. Etzel[5], T. M. Evans[3], J. Feicht[5], P. Fulda[27], M. Fyffe[3], J. A. Giaime[3,7], K. D. Giardina[3], P. Godwin[28], E. Goetz[7,16], S. Gras[1], C. Gray[4], R. Gray[24], A. C. Green[27], Anchal Gupta[5], E. K. Gustafson[5], R. Gustafson[29], J. Hanks[4], J. Hanson[3], T. Hardwick[7], R. K. Hasskew[3], M. C. Heintze[3], A. F. Helmling-Cornell[12], N. A. Holland[2], J. D. Jones[4], S. Kandhasamy[30], S. Karki[12], M. Kasprzack[5], K. Kawabe[4], P. J. King[4], J. S. Kissel[4], Rahul Kumar[4], M. Landry[4], B. B. Lane[1], B. Lantz[15], M. Laxen[3], Y. K. Lecoeuche[4], J. Leviton[29], J. Liu[17,18], M. Lormand[3], A. P. Lundgren[31], R. Macas[25], M. MacInnis[1], D. M. Macleod[25], G. L. Mansell[1,4], S. Márka[10], Z. Márka[10], D. V. Martynov[22], K. Mason[1], T. J. Massinger[1], R. McCarthy[4], S. McCormick[3], J. McIver[5], G. Mendell[4], K. Merfeld[12], E. L. Merilh[4], F. Meylahn[17,18], T. Mistry[32], R. Mittleman[1], G. Moreno[4], C. M. Mow-Lowry[22], S. Mozzon[31], T. J. N. Nelson[3], P. Nguyen[12], L. K. Nuttall[31], J. Oberling[4], Richard J. Oram[3], C. Osthelder[5], D. J. Ottaway[19], H. Overmier[3], J. R. Palamos[12], W. Parker[3,33], E. Payne[34], A. Pele[3], C. J. Perez[4], M. Pirello[4], H. Radkins[4], K. E. Ramirez[35], J. W. Richardson[5], K. Riles[29], N. A. Robertson[5,24], J. G. Rollins[5], C. L. Romel[4], J. H. Romie[3], M. P. Ross[36], K. Ryan[4], T. Sadecki[4], E. J. Sanchez[5], L. E. Sanchez[5], T. R. Saravanan[30], R. L. Savage[4], D. Schaetzl[5], R. M. S. Schofield[12], E. Schwartz[3], D. Sellers[3], T. Shaffer[4], J. R. Smith[9], S. Soni[7], B. Sorazu[24], A. P. Spencer[24], K. A. Strain[24], L. Sun[5], M. J. Szczepańczyk[27], M. Thomas[3], P. Thomas[4], K. A. Thorne[3], K. Toland[24], C. I. Torrie[5], G. Traylor[3], A. L. Urban[7], G. Vajente[5], G. Valdes[7], D. C. Vander-Hyde[13], P. J. Veitch[19], K. Venkateswara[36], G. Venugopalan[5], A. D. Viets[37], T. Vo[13], C. Vorvick[4], M. Wade[38], R. L. Ward[2], J. Warner[4], B. Weaver[4], R. Weiss[1], B. Willke[17,18], C. C. Wipf[5], L. Xiao[5], H. Yamamoto[5], Hang Yu[1], L. Zhang[5], M. E. Zucker[1,5] & J. Zweizig[5]

[3]LIGO Livingston Observatory, Livingston, LA, USA. [4]LIGO Hanford Observatory, Richland, WA, USA. [5]LIGO, California Institute of Technology, Pasadena, CA, USA. [6]Caltech CaRT, Pasadena, CA, USA. [7]Louisiana State University, Baton Rouge, LA, USA. [8]Universität Hamburg, Hamburg, Germany. [9]California State University Fullerton, Fullerton, CA, USA. [10]Columbia University, New York, NY, USA. [11]Christopher Newport University, Newport News, VA, USA. [12]University of Oregon, Eugene, OR, USA. [13]Syracuse University, Syracuse, NY, USA. [14]University of Minnesota, Minneapolis, MN, USA. [15]Stanford University, Stanford, CA, USA. [16]Missouri University of Science and Technology, Rolla, MO, USA. [17]Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Hanover, Germany. [18]Leibniz Universität Hannover, Hanover, Germany. [19]OzGrav, University of Adelaide, Adelaide, South Australia, Australia. [20]RESCEU, University of Tokyo, Tokyo, Japan. [21]OzGrav, University of Western Australia, Crawley, Western Australia, Australia. [22]University of Birmingham, Birmingham, UK. [23]Universitat de les Illes Balears, IAC3—IEEC, Palma de Mallorca, Spain. [24]SUPA, University of Glasgow, Glasgow, UK. [25]Cardiff University, Cardiff, UK. [26]The University of Mississippi, University, MS, USA. [27]University of Florida, Gainesville, FL, USA. [28]The Pennsylvania State University, University Park, PA, USA. [29]University of Michigan, Ann Arbor, MI, USA. [30]Inter-University Centre for Astronomy and Astrophysics, Pune, India. [31]University of Portsmouth, Portsmouth, UK. [32]The University of Sheffield, Sheffield, UK. [33]Southern University and A&M College, Baton Rouge, LA, USA. [34]OzGrav, School of Physics and Astronomy, Monash University, Clayton, Victoria, Australia. [35]The University of Texas Rio Grande Valley, Brownsville, TX, USA. [36]University of Washington, Seattle, WA, USA. [37]Concordia University Wisconsin, Mequon, WI, USA. [38]Kenyon College, Gambier, OH, USA.

Methods

Extended interferometer model

The model curves presented in Figs. 2, 3 and Extended Data Figs. 1, 2 are calculated from the full coupled-cavity equations of ref. [6], which are exact and omit only effects from high-order transverse optical modes. The model provided by equations (1)–(5) represents an ideal interferometer with all cavities on resonance and no optical losses. Here we extend the model to consider the dominant experimental deviations from the ideal case, without the complexity of the exact equations. This extension includes imperfect input and output efficiency, as well as the additional frequency-dependent effect on the squeezing angle from the small, unintended phase shift within the signal-recycling cavity. For the parameters used in this study, the following model is accurate to 5% or better of the exact-model quantum power spectral density at frequencies between 10 Hz and 100 Hz.

The input and output efficiency of the interferometer are introduced using two new parameters, η_i and η_o, respectively. The input efficiency represents the total fractional coupling of optical power between the squeezer cavity and the interferometer, and the output efficiency is that from the interferometer to the GW readout. They must be considered separately owing to differences in their interaction with the QRPN, leading to the expressions:

$$\Delta x^2(\Omega) = S^*[1 + \eta_o \mathcal{K}^2(\Omega)] \frac{\hbar c}{\eta_o 8k\, |G(\Omega)|^2 P_{\mathrm{arm}}} \tag{6}$$

$$1 - \eta_e = (1 - \eta_i) + \frac{1}{1 + \eta_o \mathcal{K}^2(\Omega)} (1 - \eta_o) \tag{7}$$

$$S^*(\Omega, \phi, \psi) = \eta_e S(\Omega, \phi, \psi) + (1 - \eta_e) \tag{8}$$

$$S(\Omega, \phi, \psi) = \mathrm{e}^{-2r} \cos^2(\phi - \theta^*) + \mathrm{e}^{2r} \sin^2(\phi - \theta^*) \tag{9}$$

$$\theta^* = \arctan[\mathcal{K}(\Omega)] + \frac{\Omega^2}{\gamma^2 + \Omega^2} \psi \tag{10}$$

External output loss does not change the dark-port-to-arm-cavity optical-field transmissivity $G(\Omega)$, but it does modify the dark-port-to-readout transmissivity, lowering the sensing function to $2kG(\Omega)\sqrt{\eta_o P_{\mathrm{arm}}}$. This leads to the η_o terms in equation (6), where the shot noise scales as $1/\eta_o$ but the QRPN term does not. The QRPN pertains to real motion, and its reduced influence on the optical quantum noise is compensated by the Δx calibration.

The frequency-dependent effective efficiency, η_e, accounts for the output loss $1 - \eta_o$ not being able to affect the real motion of the masses owing to radiation pressure, while the squeezed state is degraded by both input and output losses. The form of equation (7) reflects the relationship of the input, output and effective losses rather than efficiencies, and it is accurate for small losses.

The total squeezing angle shift due to the signal-recycling cavity is encoded in the parameter ψ. It appears alongside the ponderomotive effect on the squeezing angle in equation (10), except it accumulates through the cavity pole transition. This formulation is accurate for a small physical round-trip phase shift (detuning), ξ, of the interferometer signal cavity. This physical detuning results in a cavity-induced squeezing phase shift of $\psi = 10.7\xi$ calculated for the LIGO Livingston mirror parameters. Notably absent from this non-ideal model, but present in ref. [6], is the interaction between radiation pressure and the signal-recycling cavity detuning, $\xi \neq 0$, which is typically labelled an 'optical spring'. This interaction is accounted for in the calibration and exact-model curves included in our plots, but at this ξ and ψ it is not significant for the analysis. We note that the above non-ideal model is accurate to 1% in the zero-detuning case $\psi = \xi = 0$ versus a 5% discrepancy when detuning is included. Whereas strong optical springs are an alternative method of achieving sub-SQL quantum noise sensitivity, this indicates that the spring contribution is weak compared to the injected squeezing.

Measurement sequence

The data shown in Fig. 2 were taken over a 5-h period on the advanced LIGO detector. To avoid variations of classical noise and calibration, the interferometer power is held constant across all measurements. To minimize the statistical error, the majority of the measurement time is spent in the two modes plotted: three 30-min 'reference' segments with the squeezer disabled, alternating with three 30-min segments with squeezing at $\phi = 35°$. Each reference segment is followed by a squeezing segment, alternating three times to establish that the classical noise contribution is constant across the total duration. The remaining time is split across nine additional segments at varying input squeezing angles, and the final segment is a fourth reference without squeezing.

The parameters describing the status of the interferometer and squeezer during the experiment are listed in Extended Data Table 1 with uncertainties. These are also the values used in the modelling of the quantum noise calculation. Immediately before the 5-h dataset, the nonlinear parameter of the squeezer was measured to calculate r. The squeezing angle is determined ultimately through a model fit, but it agrees with our knowledge of the nonlinear conversion from the demodulation angle of the coherent control field to the observed squeezing angle and the settings during the shot-noise squeezing ($\phi = 0°$) and antisqueezing ($\phi = 90°$) datasets. The frequency-dependent contributions of the squeezing and arm-power modelling uncertainties are shown in Extended Data Fig. 1, and they do not strongly influence the model at the sub-SQL dip.

Extended Data figures

Extended Data Fig. 1 shows a variation of Fig. 2 spanning a wider frequency range. The figure includes the frequency-dependent uncertainties of equation (12) in its model curves and subtracted quantum noise plots.

Extended Data Fig. 2 shows a measurement (top) and a model (bottom) of the squeezing term $S^*(\Omega, \phi, \psi)$ of the augmented model. The quantum noise spectrum at ten additional ϕ values is determined by subtracting the classical noise contribution (previously established through the reference measurement) from the measured displacement spectrum at each ϕ. Each inferred quantum noise spectrum is then divided by the modelled quantum noise spectrum without injected squeezing (blue trace in Fig. 2) to obtain the observed squeezing term $S^*(\Omega, \phi, \psi)$. The dashed lines indicate cross-sections in other figures: the green line corresponds to $\phi = 35°$ in Fig. 2, and the magenta, navy and brown lines to the angles $\phi = 7°$, 24° and 46° shown in Fig. 3.

Uncertainty analysis for subtraction

Figure 2 shows that quantum noise accounts for only 28% of the total interferometer noise power at 40 Hz. For this reason, classical noises must be subtracted to reveal the quantum noise-limited displacement sensitivity. The interferometer is a complex instrument with such environmental sensitivity that the following considerations must be addressed to validate the subtraction. First, the fiducial quantum noise model of the reference dataset and the parameters that it relies on must be established, and the data must be calibrated. Second, the classical noise established for the reference operating mode must be representative of the classical noise during the squeezing operation. In particular, the classical noise during the reference period must not be higher than that during squeezing, which would bias our inference to underestimate the quantum noise contribution during squeezing.

To describe how uncertainty propagates through the subtraction in our measured quantum noise curves, we consider uncertainties in four sources: (a) the calibration, (b) quantum noise models, (c) statistical noise and (d) non-stationary changes in the noise contributions. D_r, D_s, M_r and M_s denote the frequency-dependent data and model spectral densities for the reference and squeezing operating cases, respectively. For our analysis, we use and plot the full coupled-cavity equations[4,6] including all losses and optical spring effects, but reiterate that the deviation between the exact model and the simplified model of equations (6)–(10) is small for our operating parameters. The post-subtraction inferred quantum noise is given as Q in the expression:

$$Q(\Omega) = D_\mathrm{s}(\Omega) - [D_\mathrm{r}(\Omega) - M_\mathrm{r}(\Omega)] \quad (11)$$

The relative error of the post-subtraction squeezed quantum noise is given by δQ and is composed of the quadrature sum of relative errors due to: the optical-sensitivity calibration, δG; the servo loop calibration, δC; the modelling uncertainty, δM_r; statistical fluctuations, δD_r and δD_s; and relative stationarity uncertainty terms, δN_t and δN_m. All of these uncertainties are frequency-dependent, but the full functional forms are omitted for brevity. These components, which are defined in the following section, contribute to the expression:

$$\begin{aligned}\delta Q^2 = \delta G^2 + \frac{1}{Q^2}[&M_\mathrm{r}^2 \delta M_\mathrm{r}^2 + (D_\mathrm{r} - D_\mathrm{s})^2 \delta C^2 \\ &+ D_\mathrm{r}^2 \delta D_\mathrm{r}^2 + D_\mathrm{s}^2 \delta D_\mathrm{s}^2 \\ &+ (D_\mathrm{r} - M_\mathrm{r})^2 (\delta N_t^2 + \delta N_m^2)].\end{aligned} \quad (12)$$

The lines of the above equation represent terms with different magnitudes of scaling terms. Given that $Q \approx M_\mathrm{s} \approx D_\mathrm{r} - D_\mathrm{s}$, the top line for the calibration and model error has terms with order-1 coefficients, indicating that the relative errors quoted in the main text remain small for the comparison to the dip model. The lower two lines of equation (12) show that the relative statistical fluctuations and stationarity uncertainties are magnified by the ratio V between the total classical power spectral density D_r and the quantum noise Q to the squeezed quantum power spectral density, which is approximately $V = 7.2$ at 40 Hz.

Calibration and modelling uncertainty

The first line of equation (12) includes the calibration and unsqueezed reference quantum noise model uncertainty terms, δG, δC and δM_r. The LIGO online calibration system determines the optical sensing function $2kG(\Omega)\sqrt{\eta_\mathrm{o} P_\mathrm{arm}}$, which affects both the model and the calibration uncertainties. To prevent double-counting in the incoherent sum, this optical gain has been isolated to the factor δG and should not be considered in δC or δM_r. The sensing function is monitored continuously by injecting displacement signals at several frequencies. Some of these appear as narrow lines in the measured spectra of Fig. 2. From these continuous injections, the bandwidth γ and the product $\eta_\mathrm{o} P_\mathrm{arm}$ are determined. In addition, parameters related to the optical spring are measured[22], but primarily affect the sensing function at frequencies <10 Hz for the measured detuning ξ, ψ. Additional lines monitor the Δx servo loop actuators to apply the frequency-dependent correction for the servo closed-loop response, which is contained in δC. The quoted frequency-dependent calibration uncertainty is the incoherent sum $\sqrt{\delta G^2 + \delta C^2}$, and error bars in Extended Data Fig. 1 include the frequency dependence.

Having factored δG out of δM_r, any error in subtracting the classical noise estimate between the reference data and the model can only arise from estimating the shot-noise and QRPN components represented by the term $g[1 + \eta_\mathrm{o}\mathcal{K}^2(\Omega)]$. Here, g is a scale factor relating the readout power to the optical field. It is unknown because the calibration system exports its sensing function in an end-to-end fashion with the photodetectors in arbitrary voltage digitization units; however, g may be well estimated using a cross-correlation method detailed below. The remaining $g\eta_\mathrm{o}\mathcal{K}^2(\Omega)$ contribution may be estimated from the factors $|G(\Omega)|^2\sqrt{P_\mathrm{arm}}$. Independent measurements establish the quoted arm power to be $P_\mathrm{arm} = 200 \pm 10$ kW; this, combined with the optical sensing gain calibration, allows us to determine the output efficiency η_o. The squeezing level at high frequencies is determined by r and $\eta_\mathrm{o}\eta_\mathrm{i}$ (see equations (7), (8)) and, using the extended datasets with $\phi = 0°$, the input efficiency η_i may be determined from the observed readout squeezing level.

The following cross-correlation method[22] is used to determine the factor g that relates the arbitrary experimental photodetector units back to the physical optical-field units. Two photodetectors are located at the readout port of the LIGO interferometer (see Fig. 1). When squeezing is not injected, the shot noise and the readout electronics noise (that is, dark noise) are uncorrelated between the two photodetectors, whereas the QRPN and all of the classical noises are correlated. If the cross-correlation and dark noise are subtracted from the total noise power for the reference dataset, then only the shot noise remains, which is calibrated to the displacement. This precisely determines the optical sensing gain in physical units, up to the uncertainty δG. The dark noise, also incoherent between the detectors, is only 1% of the shot-noise power and so contributes negligibly to the uncertainty in this subtraction.

Statistical uncertainty

The statistical uncertainty arises because the fluctuations that are intrinsic to noise also limit our ability to estimate it. With a total measurement time of T_i for a given dataset i, and a bin width of $\Delta F = 0.5$ Hz in the spectral density calculation, the relative statistical uncertainty of the inferred quantum noise power is $\delta D_i = (ET_i\Delta F)^{-1/2}$, with E the statistical efficiency accounting for the spectral estimation method. For the median method detailed below, we determine through numerical experiments on white noise that $E = 1.0$ for single-bin error bars. The bin–bin covariance due to the apodization window results in $E = 60\%$ when averaging multiple adjacent data points. The total statistical uncertainty of 8% includes both datasets δD_r and δD_s and their scaling by V in equation (12).

Measuring noise stationarity

Here we describe and characterize the terms δN_t, δN_m in the uncertainty budget of equation (12). We define these terms together as the stationarity uncertainty, and they are intended to quantify potential variations between the classical noise power, as estimated from the unsqueezed reference dataset, and the classical noise power that is actually present in the squeezing measurements. Under the presupposition that the models M_r and M_s are perfect and the statistical noise is small, these uncertainties are defined as the relative difference $D_\mathrm{s} - M_\mathrm{s} \equiv (D_\mathrm{r} - M_\mathrm{r})(1 + \delta N_t + \delta N_m)$. The two terms are distinguished as the changes to the classical noise that arise from variations in time, δN_t, and from switching the physical operating mode between the reference and squeezing, δN_m.

Stationarity uncertainty mitigation

The time variation contribution to non-stationarity, δN_t, is mitigated both through the spectral density estimation method and the use of three alternating segments for the reference and squeezed data. The aim of the alternating segments is for the operating mode to switch on a timescale shorter than the environmental variation. The environmental timescale is not known or even well defined, so instead the discontiguous segments of reference time are compared, setting a limit to the non-stationarity of the squeezing segment between them. This is done likewise for the squeezing segments surrounding a reference segment. We define a metric for the relative non-stationarity between two such discontiguous segments to be:

$$\mathcal{N}_{ij} = 2\frac{D_i - D_j}{D_i + D_j}. \tag{13}$$

Each pair of datasets is used to make an estimate of the noise contribution varying at and below the separation timescale of the datasets; here, 1 h. The estimates $\mathcal{N}_{ij}$ are limited by the statistical error of the constituent reference and squeezing datasets, denoted as $\mathcal{N}_{\mathrm{R}ij}$ and $\mathcal{N}_{\mathrm{S}ij}$, respectively, and they are shown in Extended Data Fig. 3. Because each pair comprises only a fraction of its full dataset, multiple estimates are combined to reduce the statistical uncertainty.

$$\mathcal{N}_{\Sigma}^2 = \frac{1}{6}(\mathcal{N}_{\mathrm{R12}}^2 + \mathcal{N}_{\mathrm{R23}}^2 + \mathcal{N}_{\mathrm{R31}}^2 + \mathcal{N}_{\mathrm{S12}}^2 + \mathcal{N}_{\mathrm{S23}}^2 + \mathcal{N}_{\mathrm{S31}}^2). \tag{14}$$

Finally, these metrics must be related to the stationarity term δN_t. The averaged non-stationary power $\mathcal{N}_{\Sigma}^2$ represents an estimate of the time-varying contribution between adjacent reference and squeezing segments, of which there are three. For many such segments, assuming random fluctuations to the environmental noise level at the alternation timescale, the contributions add in quadrature to give $\delta N_t^2 \lesssim \mathcal{N}_{\Sigma}^2/3$. We then propagate the statistical noise limits for segments with one-third of the length of the total reference time T. This arrives at the statistical limit to our stationarity uncertainty of $\delta N_t \approx \sqrt{2}\,(ET\Delta F)^{-1/2}$. Because the total squeezing data time is also T, our limit to the time variation contribution to non-stationarity is evaluated to be the same as the total statistical uncertainty from both the squeezed and unsqueezed datasets, $\delta N_t^2 \approx \delta D_\mathrm{r}^2 + \delta D_\mathrm{s}^2$. In addition to the individual pairs, Extended Data Fig. 3 shows the combined estimate $\mathcal{N}_{\Sigma}^2$.

The operating mode variation component δN_m of non-stationary noise is constrained by the following considerations. The first is that it is quantitatively constrained by the data at the additional squeezing angles depicted in Fig. 3 and Extended Data Fig. 2. There, the same classical noise estimate is subtracted and the model curves maintain their agreement with the inferred quantum noise at alternate squeezing angles. Those datasets, however, have limited statistical bounds owing to their short duration. The term δN_m may be considered small for the following physical reasons. The primary reason is that during the time without squeezing, the optical path is not changed, and only the optical parametric oscillator cavity[19] which produces the squeezed states is operated off resonance to stop its nonlinear parametric interaction. This means that environmental scatter noise—the very-low-power light leaking from the interferometer to the squeezer system—does not impinge on different scattering surfaces between the two modes. In the event that such scatter does matter, the fourth reference taken at the end of the entire measurement period uses an in-vacuum beam diverter to block the path to the squeezer. Testing that fourth reference against the other three through the $\mathcal{N}_{ij}$ method shows no substantial changes to the classical noise.

If the classical noise does change from the switch to squeezing, we argue that the addition of the nonlinear parametric interaction from the squeezer to this scattered light is more likely to increase the noise only during the squeezing segments. This implies that the measurement should not be biased low and will not overestimate how much we have surpassed the SQL. Indeed, the few data points in Fig. 2 and Extended Data Fig. 1 that exceed the model beyond the statistical fluctuations may be due to such a squeezer-specific noise source. We attribute the minimal classical noise contribution to the use of a travelling-wave optical parametric oscillator cavity, the in-vacuum suspended layout and coherent control implementation[19].

Spectral density estimation

To mitigate the non-stationary noise power contributions, a statistically robust median-based computation method is used to calculate the sampled power spectral density. Based on the above considerations, we claim that the classical noise is established to be stationary in these datasets; however, it is known from astrophysical analysis that these complex detectors have intermittent time-resolved glitches and artefacts of varying strength. Intervals of excess noise are nontrivial to identify owing to the inherently random nature of noise, and time-resolved noise power vetoes can introduce selection bias. We use the Welch–Bartlett overlap method to estimate the power spectral density with no selection vetoes. Instead, rather than averaging the individual spectra independently at each frequency, the sample median at each frequency is taken. This generates a bin-by-bin median strain spectral power density.

Initially, the entire period for a given spectral density estimate is split into N 2-s segments, where each segment overlaps the segment before it by 50%, implementing the Welch method. For each segment, the time series is linearly detrended and a Hann window is applied; then, the time series is converted to a displacement spectrum using a Fourier transform. The collection of segments gives N estimates of the power density in each frequency bin, each of which nominally follows a χ^2 distribution on two variables (the real and imaginary parts of the Fourier transform), but the distribution has an extended tail due to glitches and transients of the detector. The median is picked for each frequency bin, and then a computed scale factor is applied to convert the distribution median to the mean noise power. This technique is unbiased for stationary noise and greatly improves the robustness to glitches and non-stationary contributions, without selection bias from time-domain band-limited noise vetoes. The downside is that the statistical efficiency is approximately $\sqrt{2}$ worse than the typical Welch method for a given spectrum-averaging time.

Data availability

Source data for Figs. 2, 3, Extended Data Figs. 1–3 and other data pertaining to this study are available from the corresponding authors upon reasonable request.

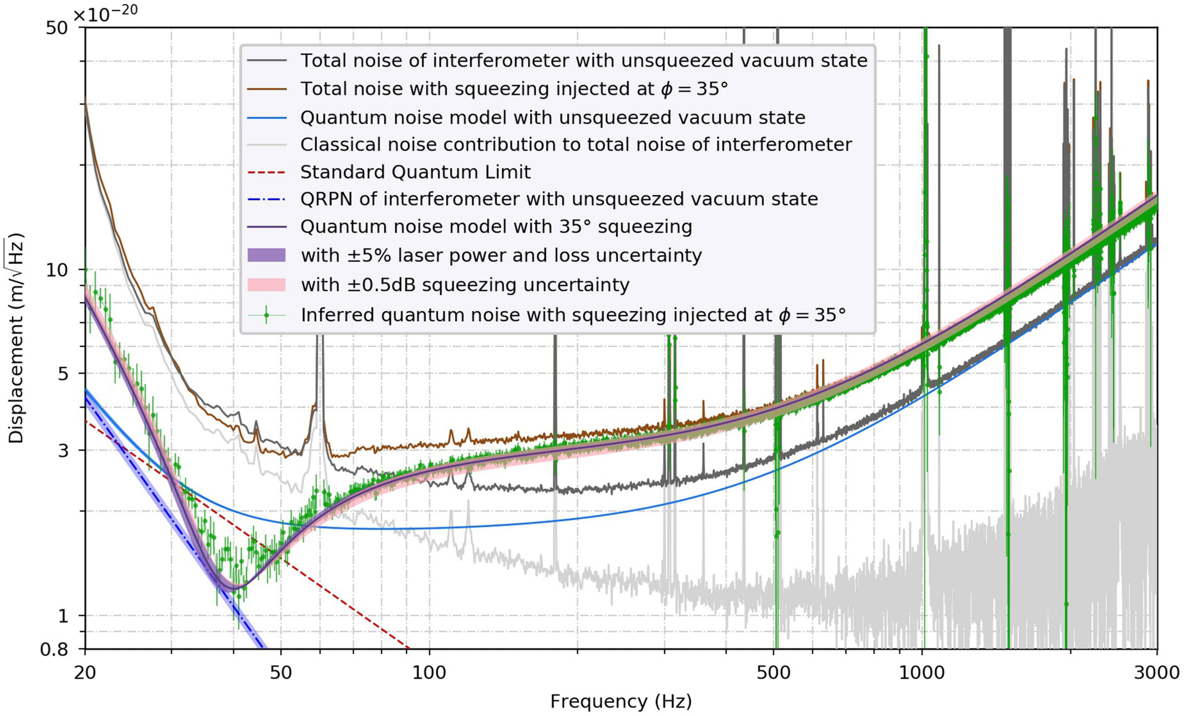
22. Kiwamu, I. Time domain implementation of dcpd cross correlation. *Technical Note* T1700131 (LIGO Virgo, 2017); https://dcc.ligo.org/LIGO-T1700131/public.

Acknowledgements LIGO was constructed by the California Institute of Technology and the Massachusetts Institute of Technology with funding from the National Science Foundation, and operates under Cooperative Agreement number PHY-1764464. Advanced LIGO was built under grant number PHY-0823459. The authors gratefully acknowledge the support of the Australian Research Council under the ARC Centre of Excellence for Gravitational Wave Discovery grant number CE170100004, Linkage Infrastructure, Equipment and Facilities grant number LE170100217 and Discovery Early Career Award number DE190100437; the National Science Foundation Graduate Research Fellowship under grant number 1122374; the Science and Technology Facilities Council of the United Kingdom; and the LIGO Scientific Collaboration Fellows programme.

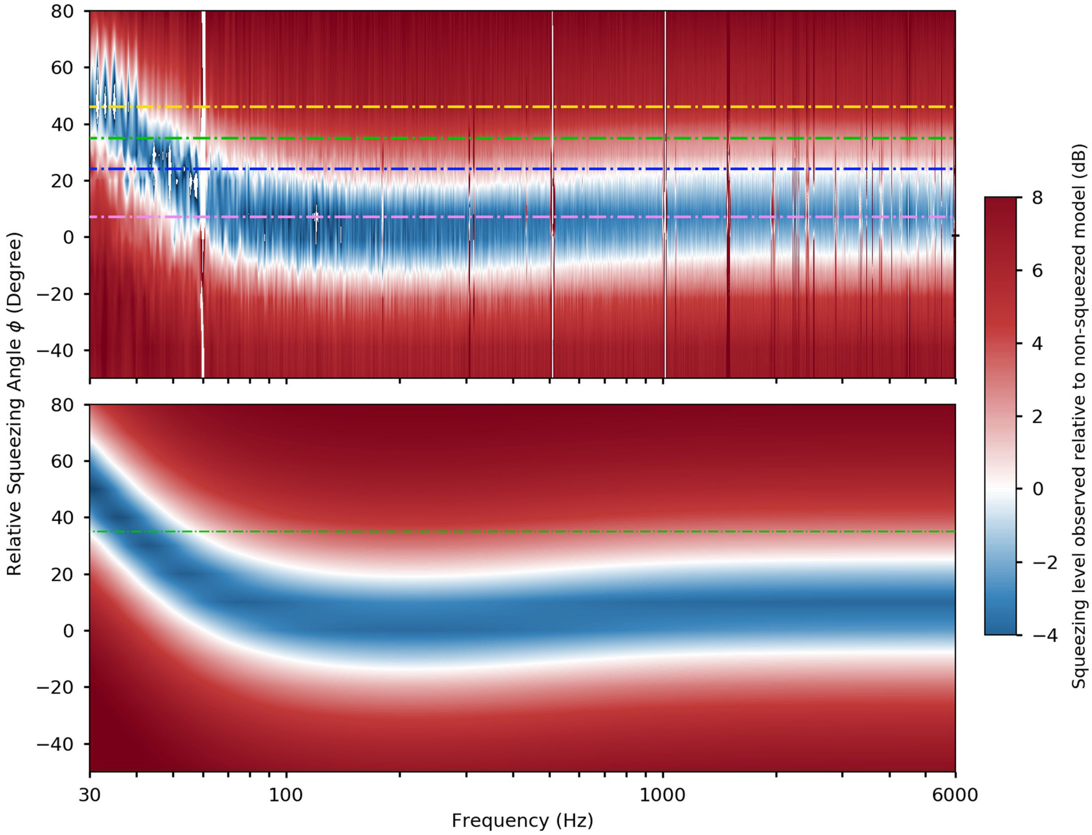
Author contributions The measurements presented in this paper were performed with the 4-km detector at the LIGO Livingston Observatory using a novel squeezed light source. Haocun Yu performed all of the measurements. Haocun Yu and L.M. carried out the analysis of the data. M. Tse, Haocun Yu and N.K. built and commissioned the squeezed-light sources. L.B. led the squeezed-light upgrade of the LIGO detectors, involving contributions from a large number of people within the LIGO Laboratory, the Australian National University and other members of the LIGO Scientific Collaboration. N.M. led the experimental campaign to measure sub-SQL quantum noise in the Advanced LIGO detector. Haocun Yu, L.M., M. Tse, L.B. and N.M. contributed directly to the preparation of the manuscript.

Competing interests The authors declare no competing interests.

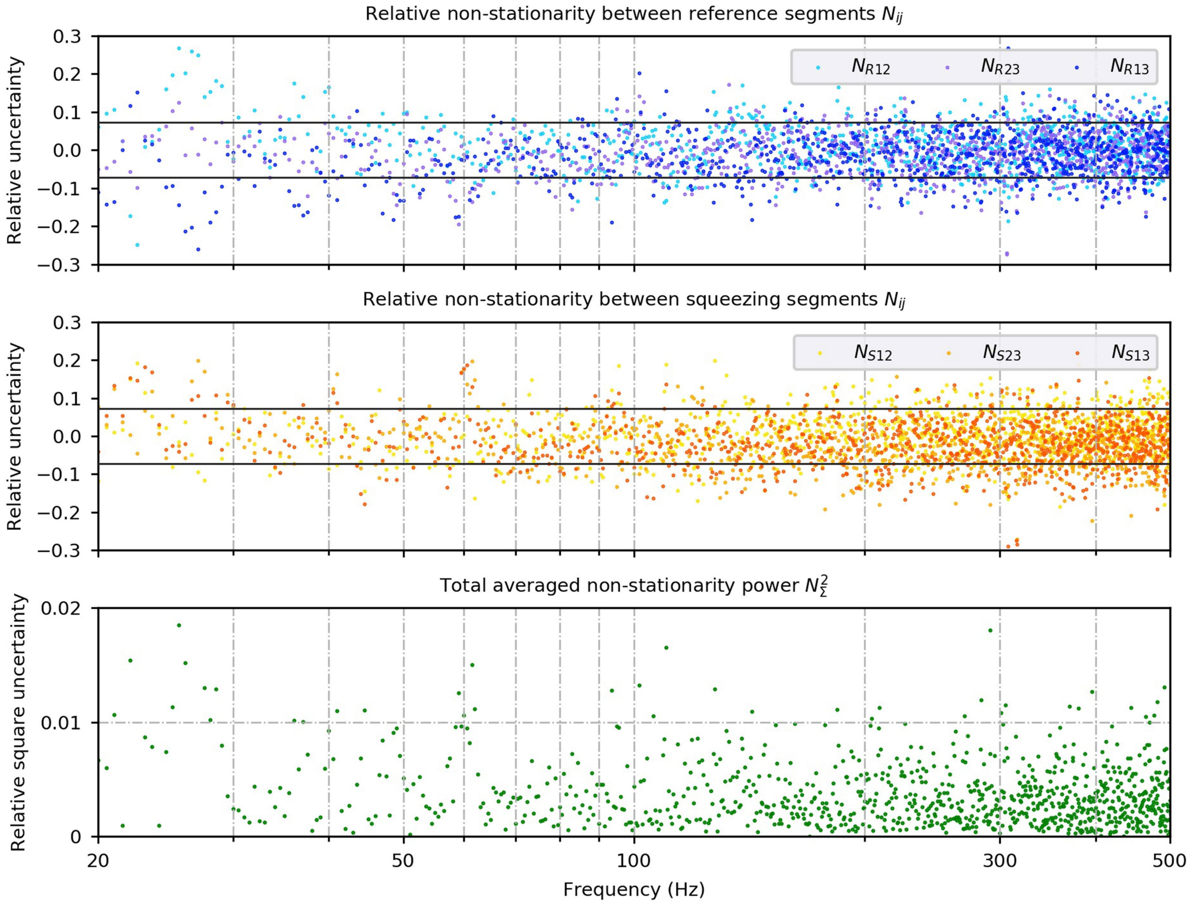
Additional information
Correspondence and requests for materials should be addressed to Haocun Yu or L.M.
Peer review information *Nature* thanks Albert Schliesser, Valeria Sequino and Kentaro Somiya for their contribution to the peer review of this work.
Reprints and permissions information is available at http://www.nature.com/reprints.



Extended Data Fig. 1 | Spectral density measurements revealing sub-SQL quantum noise of the interferometer with uncertainties. The black and brown traces show the measured total noise level of the interferometer with the unsqueezed vacuum state (the reference) and injected squeezing at 35°, respectively. The grey curve shows the classical noise contribution to the total noise of the interferometer, which is independent of the squeezer state. The solid blue curve shows the quantum noise model and includes the 5% uncertainty in the arm power, compensated by the output optical loss to maintain the calibrated sensing function. The inferred quantum noise (green curve) and error bars include all uncertainty terms present in equation (12), as estimated in Methods, including the frequency dependence. The quantum noise model with 35° squeezing (purple line) is shown with the 5% arm power uncertainty (purple shading) and the 0.5-dB uncertainty of the squeezing generated by the squeezer (pink shading). The free-mass SQL is shown by the dashed red line, and the pure QRPN contribution of the interferometer with the unsqueezed vacuum state is shown by the dashed blue line and includes the uncertainty in the arm power.



Extended Data Fig. 2 | Squeezing level of the interferometer over the full range of squeezing angles. Contour plot of squeezing level $S^*(\phi, \theta, \psi)$ detected in the interferometer as a function of the frequency and squeezing angle ϕ (top) and the corresponding theoretical model (bottom). The dashed lines indicate cross-sections in other figures. The green dashed line shows $\phi = 35°$ in Fig. 2, and the magenta, navy and orange lines correspond to the angles shown in Fig. 3.



Extended Data Fig. 3 | Individual and combined estimates of non-stationary noise between measurement segments. The two upper plots show the relative time variation of noise between each pair of reference and squeezing measurement segments, respectively. The black lines show 2σ or a 95% confidence level. The bottom plot shows the combined non-stationary power defined by equation (14).

Extended Data Table 1 | Interferometer and squeezer parameters used for modelling the Advanced LIGO detector in Livingston

Interferometer Parameter	Value
Laser power in the arm cavity (P_{arm})	200±10 kW
Optical loss before interferometer ($1-\eta_i$)	17.2%
Optical loss after interferometer ($1-\eta_o$)	17.4%
SRM phase detuning (ξ)	15 mrad

Squeezer Parameter	Value
Measured OPO nonlinear gain	4.4±0.1
Squeezing ideally generated by OPO (e^{-2r})	9.8±0.15 dB
Squeezer phase noise ($\delta\phi$)	0-50 mrad
Squeezing quadrature rotation angle (ϕ)	35°
Max phase squeezing in interferometer	3.3 dB
Max phase anti-squeezing in interferometer	7.7 dB

Fluorescent bicolour sensor for low-background neutrinoless double β decay experiments

https://doi.org/10.1038/s41586-020-2431-5

Received: 15 September 2019

Accepted: 3 April 2020

Iván Rivilla[1], Borja Aparicio[2], Juan M. Bueno[3], David Casanova[1,4], Claire Tonnelé[1], Zoraida Freixa[4,5], Pablo Herrero[1], Celia Rogero[1,6], José I. Miranda[7], Rosa M. Martínez-Ojeda[3], Francesc Monrabal[1,4], Beñat Olave[8], Thomas Schäfer[4,8], Pablo Artal[3], David Nygren[9], Fernando P. Cossío[1,2] ✉ & Juan J. Gómez-Cadenas[1,4] ✉

Observation of the neutrinoless double β decay is the only practical way to establish that neutrinos are their own antiparticles[1]. Because of the small masses of neutrinos, the lifetime of neutrinoless double β decay is expected to be at least ten orders of magnitude greater than the typical lifetimes of natural radioactive chains, which can mimic the experimental signature of neutrinoless double β decay[2]. The most robust identification of neutrinoless double β decay requires the definition of a signature signal—such as the observation of the daughter atom in the decay—that cannot be generated by radioactive backgrounds, as well as excellent energy resolution. In particular, the neutrinoless double β decay of ^{136}Xe could be established by detecting the daughter atom, ^{136}Ba^{2+}, in its doubly ionized state[3–8]. Here we demonstrate an important step towards a 'barium-tagging' experiment, which identifies double β decay through the detection of a single Ba^{2+} ion. We propose a fluorescent bicolour indicator as the core of a sensor that can detect single Ba^{2+} ions in a high-pressure xenon gas detector. In a sensor made of a monolayer of such indicators, the Ba^{2+} dication would be captured by one of the molecules and generate a Ba^{2+}-coordinated species with distinct photophysical properties. The presence of such a single Ba^{2+}-coordinated indicator would be revealed by its response to repeated interrogation with a laser system, enabling the development of a sensor able to detect single Ba^{2+} ions in high-pressure xenon gas detectors for barium-tagging experiments.

Double β decay ($\beta\beta$) is a very rare nuclear transition in which a nucleus with Z protons decays into a nucleus with $Z+2$ protons and the same mass number A. The decay can occur only if the initial nucleus is less strongly bound than the final nucleus, and both of them are more strongly bound than the intermediate $Z+1$ nucleus. Two decay modes are usually considered: (i) The standard two-neutrino mode ($\beta\beta2\nu$), consisting of two simultaneous β decays, $(Z,A)\rightarrow(Z+2,A)+2e^-+2\bar{\nu}_e$ (e^-, electron; $\bar{\nu}_e$, electron antineutrino), which has been observed in several isotopes with typical half-lives in the range 10^{18}–10^{21} yr; and (ii) The neutrinoless mode ($\beta\beta0\nu$), $(Z,A)\rightarrow(Z+2,A)+2e^-$, which violates lepton-number conservation and can occur if and only if neutrinos are Majorana particles[1]—that is, identical to their antiparticles. An unambiguous observation of such a decay would have deep implications in particle physics and cosmology, offering a mechanism for leptogenesis[9] and a potential explanation for the cosmic asymmetry between matter and antimatter[10]. Furthermore, Majorana neutrinos could provide an explanation of the smallness of the neutrino mass compared with those of other leptons, through the so-called see-saw mechanism[11–13].

Double β decay ($\beta\beta$) experiments have been searching for $\beta\beta0\nu$ in several isotopes for more than half a century, without finding clear evidence of a signal so far. The current best lower limit on the lifetime ($T^{0\nu}_{1/2}$) of the $\beta\beta0\nu$ processes has been obtained for the isotope ^{136}Xe, for which $T^{0\nu}_{1/2} > 10^{26}$ yr (ref. 14). Two other isotopes, ^{76}Ge and ^{130}Te, have also been studied with similar sensitivities, yielding no evidence of $\beta\beta0\nu$ decay[15,16]. A new generation of $\beta\beta0\nu$ experiments will aim to improve the sensitivity to $T^{0\nu}_{1/2}$ by at least one, and eventually two, orders of magnitude[17]. These searchers will require very large exposures, measured in ton-years, but even more importantly, a greatly enhanced capability to suppress backgrounds from false events. The most obvious background to $\beta\beta0\nu$ is the $\beta\beta2\nu$ decay, which also produces two electrons and the same daughter atom as the neutrinoless mode while having a

---

[1]Donostia International Physics Center (DIPC), San Sebastián/Donostia, Spain. [2]Department of Organic Chemistry I, University of the Basque Country (UPV/EHU), Centro de Innovación en Química Avanzada (ORFEO-CINQA), San Sebastián/Donostia, Spain. [3]Laboratorio de Óptica (LOUM) & Centro de Investigación en Óptica y Nanofísica (CiOyN), University of Murcia, Murcia, Spain. [4]Ikerbasque, Basque Foundation for Science, Bilbao, Spain. [5]Department of Applied Chemistry, Faculty of Chemistry, University of the Basque Country (UPV/EHU), San Sebastián/Donostia, Spain. [6]Materials Physics Center CFM (CSIC-UPV/EHU), San Sebastián/Donostia, Spain. [7]SGIker NMR Facility, University of the Basque Country (UPV/EHU), San Sebastián/Donostia, Spain. [8]NanoBioSeparations Group, POLYMAT, University of the Basque Country (UPV/EHU), San Sebastián/Donostia, Spain. [9]Department of Physics, University of Texas at Arlington, Arlington, TX, USA. ✉e-mail: fp.cossio@ehu.es; jjgomezcadenas@dipc.org

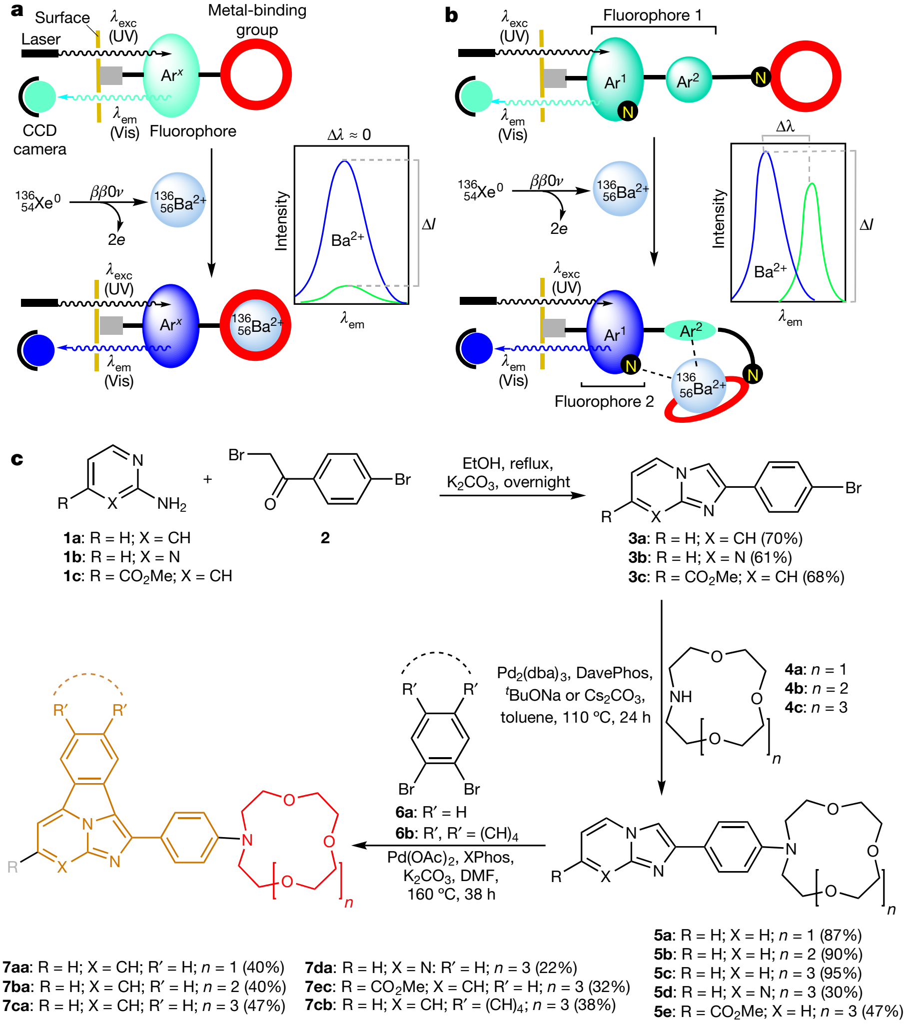


Fig. 1 | Design and synthesis of a family of FBIs. a, Components of a fluorescent monocolour indicator. UV, ultraviolet; Vis, visible; CCD, charge-coupled device; em, emission; exc, excitation; Ar^x, aryl or heteroaryl group (monocyclic or polycyclic). **b**, Components of an FBI analogue, showing the coupling–decoupling between the fluorophore and the metal-binding group. The respective expected fluorescent emission spectra are also shown. The blue and green lines in the graphs in **a**, **b** represent the emission spectra in chelated and unchelated indicators, respectively. The different behaviour of the two types of fluorescent indicator is shown: in a monocolour indicator, the increase in emission intensity after cation complexation (ΔI) is produced at the same wavelength as the unchelated fluorophore ($\Delta\lambda \approx 0$), whereas in the case of an FBI, the difference ΔI is produced at a different wavelength ($\Delta\lambda \neq 0$). The possible participation of nitrogen heteroatoms and the rotation of one aryl group (Ar^2) are also highlighted. **c**, Chemical synthesis of a family of FBIs. The synthetic route starts from pyridines (or pyrimidines) and **1a**–**1c**, 4-bromoacetophenone **2**, to form adducts **3a**–**3c**. Coupling of these latter intermediates with aza-crown ethers **4a-c** yields compounds **5a**–**5c**, which reacts with 1,2-dibromoarenes **6a**,**b** to give the FIB candidates **7aa**–**7cb**. Numbers in parentheses correspond to the chemical yields (average values after three or five independent experiments) of isolated pure products. DavePhos, 2'-(dicyclohexylphosphino)-*N*,*N*-dimethyl-2-biphenylamine; dba, dibenzylideneacetone; XPhos, dicyclohexyl(2',4',6'-triisopropyl-2-biphenylyl) phosphine.

much faster decay rate. Near the end energy (Q), however, the $\beta\beta2\nu$ process is very strongly suppressed by kinematics, and its contamination to the $\beta\beta0\nu$ signal is very small for a detector with good energy resolution[18].

Instead, owing to the irreducible presence of trace amounts of the radioactive decays chains of ^{238}U and ^{232}Th in the materials of the detector, the corresponding false signatures need to be suppressed by a very large factor. The decays of other radioactive isotopes created by neutron activation are also a concern. All $\beta\beta$ experiments are built with ultrapure materials, operate in underground laboratories (to mitigate the impact of cosmic rays) and are protected by massive, ultrapure shields. These strategies reduce the ambient background by many orders of magnitude, but putative $\beta\beta0\nu$ events must still be extracted against tens of millions of spurious interactions.

The most powerful discriminant against backgrounds other than $\beta\beta2\nu$ would be the detection of the daughter atom, which is displaced by two steps in the periodic table relative to its parent. In particular, the decay $^{136}\text{Xe} \rightarrow {}^{136}\text{Ba}^{2+} + 2e^- + (2\bar{\nu}_e)$ will create a Ba^{2+} dication as the most likely outcome in xenon gas. In pure xenon gas, no known radioactive process will produce this ion in coincidence with two electrons. The implementation of a robust Ba^{2+} detection technique would facilitate the positive identification of a $\beta\beta0\nu$ candidate. The possibility of barium tagging in a xenon time-projection chamber (TPC) was proposed in 1991 by Moe[3] and has been extensively investigated for the past two decades[4,19,20].

Recently the nEXO collaboration demonstrated the imaging and counting of individual barium atoms in solid xenon by scanning a focused laser across a solid xenon matrix deposited on a sapphire window[5]. This is a promising step towards barium tagging in liquid xenon. The technique originally proposed by Moe and being pursued by nEXO relies on Ba^{+} fluorescence imaging using two atomic excitation levels in very-low-density gas. In liquid xenon, recombination is frequent

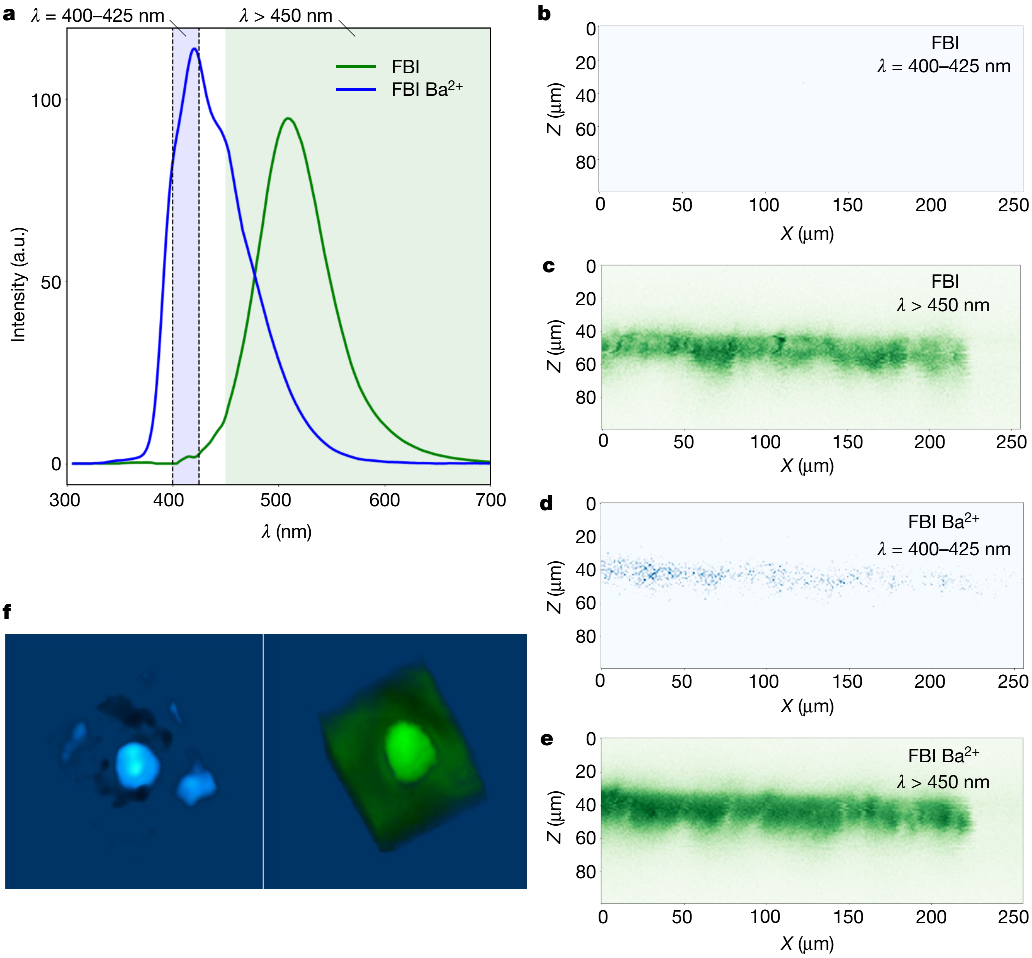


Fig. 2 | Response of the FBI. a, Emission spectrum of the SF (green line) and SBF (blue line) samples after silica subtraction (the SF spectrum is scaled by a factor of C_r with respect to the SBF spectrum). **b**, *Z–X* profile of the control pellet, SFpA, showing no signal in the deep-blue region (400, 425) nm, where the contribution from unchelated molecules is negligible. **c**, *Z–X* profile of SFpA in the green region ($\lambda > 450$ nm), showing intense green emission from the unchelated molecules. **d**, *Z–X* profile of the sublimated pellet, SFpB, showing a clear signal in the deep-blue region (400, 425) nm due to the molecules chelated by the barium perchlorate. **e**, *Z–X* profile of SFpB in the green region ($\lambda > 450$ nm), showing intense green emission from both chelated and unchelated molecules. **f**, 3D tomography images of SFpB, obtained with our TPA microscopy setup, passed through the blue and the green filters. The images reveal the shape of a tiny section (a square of 75 μm^2 size), showing the same landscape for both chelated and unchelated molecules.

and the barium daughters are distributed across charge states from 0 to 2+ (ref. [21]), with sizeable populations of neutral Ba and Ba^{+}. In the high-pressure gas phase, however, the initially highly ionized barium daughter quickly captures electrons from neutral xenon, stopping at Ba^{2+}, beyond which recombination is minimal[22].

A molecule with a response to optical stimulation that changes when it forms a supramolecular complex with a specific ion is a fluorescent indicator, and ions thus non-covalently bound to molecules are generally referred to as being chelated. In 2015, Nygren proposed a Ba^{2+} sensor based on fluorescent molecular indicators that could be incorporated within a high-pressure gas xenon TPC (HPXe)[6], such as those being developed by the NEXT Collaboration[17,23–25]. The concept was further developed in ref. [7] and was followed by an initial proof-of-concept study[8], which resolved individual Ba^{2+} ions on a thin quartz plate with Fluo-3 (a common indicator in biochemistry) suspended in polyvinyl alcohol (PVA) to immobilize the molecular complex and facilitate optical imaging. The experiment demonstrated single-ion sensitivity (with a root-mean-square super-resolution of 2 nm), which was confirmed by single-step photobleaching, and provided an essential step towards barium tagging in an HPXe.

However, an experiment aiming to detect Ba^{2+} in an HPXe requires a sensor that differs substantially from that used in ref. [8]. First, the surface density of indicators in the sensor needs to be high to ensure maximum ion capture efficiency. Second, the indicators must be able to form a supramolecular complex with Ba^{2+} in a dry medium, that is, the Gibbs energy of the process in xenon gas must be negative. Third, the indicators must respond to optical stimulation with a very distinctive signal that allows unambiguous identification of the molecule that has chelated the single ion produced in the $\beta\beta0\nu$ decay and good discrimination from the background due to the uncomplexed molecules in the surroundings. In other words, the discrimination factor, *F*, between the response (in a dry medium) of the chelated indicator and the residual response of unchelated molecules must be large. A considerable step in developing dry sensors was carried out in ref. [26], where molecular compounds based on aza-crown ethers and using fluorophores such as pyrene[27,28] and anthracene[29] were studied.

In this paper we demonstrate an important step towards a barium-tagging experiment in an HPXe, using a fluorescent bicolour indicator (FBI) as the core of a sensor that detects single Ba^{2+} ions in a high-pressure gas detector. The indicator is designed to bind strongly to Ba^{2+} and to shine very brightly when complexed with Ba^{2+}. Furthermore, the emission spectrum of the chelated indicator is considerably blue-shifted with respect to the unchelated species, allowing an additional discrimination of almost two orders of magnitude.

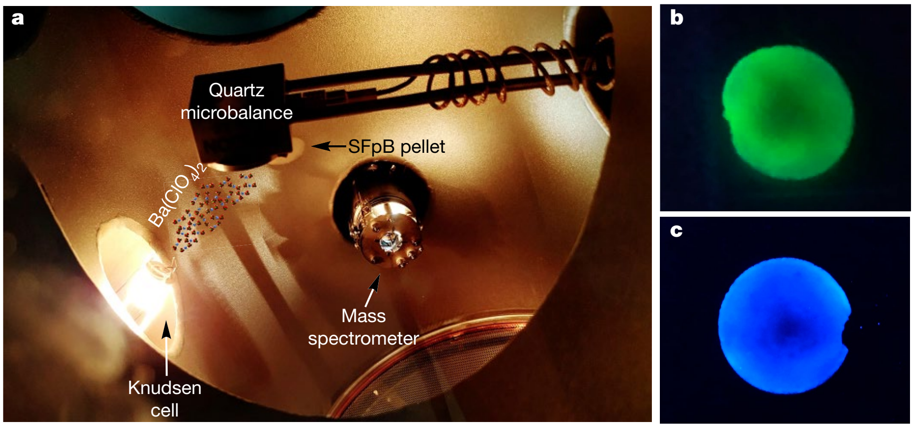


Fig. 3 | Sublimation of $Ba(ClO_4)_2$ on the FBI. a, Experimental setup. Photograph of the interior of the UHV chamber used for sublimation. The positions of the pellet, evaporator, quartz microbalance and mass spectrometer are indicated. **b**, **c**, Photographs of the pellet before (**b**) and after (**c**) the sublimation. In both cases, the excitation light is 365 nm. We note the characteristic green colour of unchelated FBI before the sublimation and the blue shift after the sublimation, which shows a large density of chelated molecules.

Design and synthesis of FBI compounds

Our criteria for designing FBIs are summarized in Fig. 1. The indicator includes, as essential components, a metal-binding group (a convenient moiety is a coronand formed by an *N*-aryl-aza-crown ether[30,31]) and a fluorophore, in line with previously developed designs for fluorescent sensors able to capture metal cations in solution[32]. Figure 1a shows the expected behaviour of a fluorescent monocolour indicator, in which the fluorophore does not modify substantially its π-molecular orbital structure upon metal coordination. In these hydrocarbon or heterocyclic scaffolds, an electron-donating group close to the fluorophore (for instance, an amino group of the aza-crown ether) can promote a photoinduced electron transfer that quenches the fluorescence in the absence of a binding cation. By contrast, sensor–cation complexation results in an off–on enhancement of the photoemission intensity[33] with $\Delta\lambda \approx 0$ (Fig. 1a). Therefore, in general only changes in the intensity of the emitted fluorescent signal upon Ba^{2+} complexation should be observed under this photoinduced electron transfer mechanism. This kind of sensor has been used in aqueous solution for metals of biological interest[34] and mainly for the capture of cations such as K^+ by using bicylic aza-cryptands[35]. Figure 1b illustrates the desired behaviour of an FBI indicator upon binding to Ba^{2+} ions. A convenient way to generate this kind of sensor with $\Delta\lambda \neq 0$ consists of generating an intramolecular photoinduced charge transfer (PCT) by modifying the interaction of an electron-donating group with the rest of the fluorophore[32]. Upon coordination with the cation, the change in the dipole moment of the supramolecular entity can generate a Stokes shift. However, in general these PCT phenomena promote only slight blue shifts[36] and depend on the polarity of the environment, thus being strongly affected by solvent effects. Actually, most PCT sensors work in water and bind cations such as Na^+ and K^+ by means of bicyclic aza-cryptands[37,38], among other groups such as acidic chelators or podands. Therefore, the design and chemical synthesis of efficient FBIs with large enough $\Delta\lambda$ values in the gas phase still constitutes an important challenge.

Within this context, we require that: (i) the chelating group binds the cation with a high binding constant; (ii) the indicator response in a dry medium is preserved and preferably enhanced with respect to the response in solution; and (iii) the fluorophore exhibits a distinct response in the visible region for the chelated and unchelated states (thus the term 'bicolour indicator'). To that end, the synthesis of FBI compounds incorporates a custom-designed fluorophore possessing two aromatic components, denoted as Ar^1 and Ar^2 in Fig. 1b that are connected by a free-rotating σ bond. The main fluorophore component Ar^1 consists of a nitrogen-containing aromatic polyheterocycle[39–42] that can bind the Ba^{2+} cation, thus modifying its electronic structure and decoupling this moiety from Ar^2, which in turn can generate a π–cation interaction[43] (Fig. 1b). The expected shift in response to the coordination should provide a strong signature of a bound indicator, exhibiting a blue shift over a background of unbound species. Furthermore, we require that the indicator response does not form supramolecular complexes with light elements in the barium column of alkaline earth elements (such as beryllium, calcium and magnesium) as well as with other close alkali ions that are frequently found in the environment, such as Na^+ and K^+.

The chemical synthesis of our sensors is shown in Fig. 1c. The process starts with the double addition–elimination reaction between 2-aminopyridines (X = CH) **1a**,**c** or 2-aminopyrimidine **1b** (X = N) and 2,4-dibromoacetophenone **2**. Bicyclic heterocycles **3a**–**3c** react with aza-crown ethers **4a-c** in the presence of a Pd(0)/DavePhos catalytic system to generate intermediates **5a**–**5e** in moderate (30%) to very good (95%) yields. Finally, these latter adducts are coupled with aromatic 1,2-dibromides **6a**,**b** by means of a catalytic system formed by a Pd(II) salt and XPhos to yield the desired FBI compounds **7aa**–**7cb**. In this latter step, the formal (8 + 2) reactions are carried out in the presence of potassium carbonate or caesium carbonate (compound **7ec**) as weak bases.

Finally, we performed experiments to determine the photo-physical properties of compounds **7**. The results of these experiments, which are described in Methods, allowed us to select compound **7ca** as the optimal combination of structural and electronic features that fulfil our design criteria. We refer henceforth to compound **7ca** as FBI.

Discrimination factor

To demonstrate the performance of our FBI as a Ba^{2+} sensor, we adopted silica gel as a solid-phase support. Adsorption of the molecule on the silica surface permits the exposure of at least one side of its crown ether moiety to the interaction with Ba^{2+} cations. In addition, this solid–gas interface topology preserves the conformational freedom required to reach the coordination pattern observed in our calculations (see Methods for further information), keeping the essential features of our design, in particular the Ba^{2+}-induced colour shift.

Two samples were manufactured. Sample SF was prepared by depositing on a silica pellet 2.3×10^{-5} mmol of FBI (from a CH_3CN solution) per milligram of silica. Sample SBF was formed by depositing 7.4×10^{-8} mmol of FBI (from a CH_3CN solution) per milligram of silica on a silica pellet saturated with barium perchlorate. The optimal concentration of barium salt was determined by a titration experiment

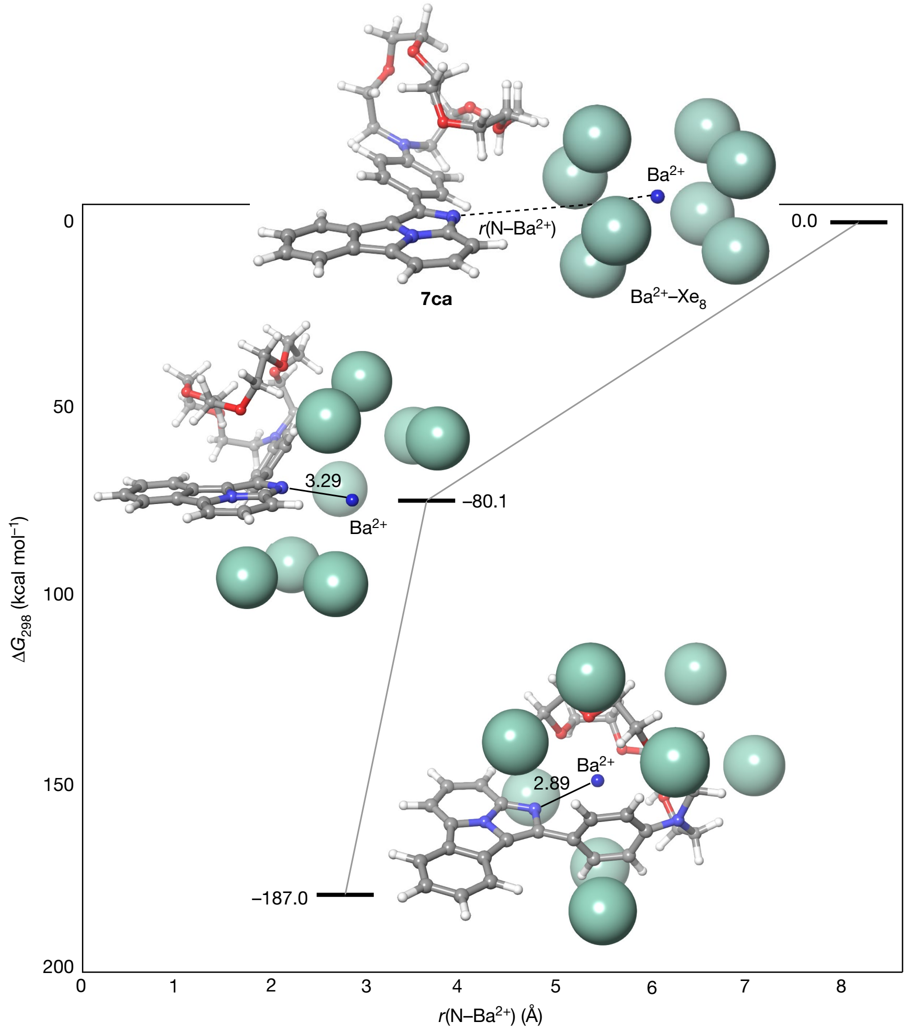


Fig. 4 | Computed structures of FBI (7ca) and a Ba^{2+}Xe$_8$ cluster at different N–Ba^{2+} distances. The geometries and energies shown were computed using DFT (see Methods for further details). Xenon atoms are represented using the Corey–Pauling–Koltun (CPK) space-filling model. The remaining atoms are represented using a ball-and-stick model and the CPK colouring code. Relative free energies (ΔG_{298}), have been computed at 25 °C (298 K).

described in Methods. The ratio between the FBI concentrations of SF and SBF was $C_r = 310 \pm 6$, where the 2% relative error was determined by propagating the uncertainties in the measurements of the volumes of the solutions. Figure 2a shows the emission spectra of the SF (SBF) samples for an excitation light of 250 nm, recorded by a fluorimeter after evaporating the solvent and subtracting the background signal due to the silica (see Methods for a discussion).

A robust separation between SF and SBF can be achieved by selecting a blue-shifted wavelength range of $\lambda_f = (\lambda_{\min}, \lambda_{\max})$ using a band filter. We call $C(\lambda)$ the emission spectrum of the chelated molecules (for example, the blue curve in Fig. 2a) and $U(\lambda)$ that of the unchelated molecules (green curve in Fig. 2a). The fraction of $C(\lambda)$ selected by the filter $f_c = c'/C$, where $c' = \int_{\lambda_{\min}}^{\lambda_{\max}} C(\lambda)\mathrm{d}\lambda$ and $C = \int C(\lambda)\mathrm{d}\lambda$. Analogously, the fraction of $U(\lambda)$ selected by the filter is $f_u = u/U$ with $u = \int_{\lambda_{\min}}^{\lambda_{\max}} U(\lambda)\mathrm{d}\lambda$ and $U = \int U(\lambda)\mathrm{d}\lambda$. By defining $D_r = f_c/f_u$, the discrimination factor is simply:

$$F = D_r C. \tag{1}$$

For this study we chose a band filter with $\lambda_f = (400, 425)$ nm, corresponding to the region shaded in blue in Fig. 2a. A larger separation could be obtained by including smaller wavelengths (for example, selecting $\lambda_f < 400$ nm), but the fluctuations associated with the subtraction of the baseline and the rapid variation of $C(\lambda)$ would also result in large uncertainties. We find $f_c = 0.29 \pm 0.03$, $f_u = 0.0036 \pm 0.0007$ and $D_r = 80 \pm 18$ (all uncertainties denote the root-mean-square deviation). The approximately 20% relative error in the estimation of f_u is dominated by the subtraction of the baseline, whereas the approximately 10% relative error in the estimation of f_c is found by varying the range of the filter by ±1 nm. Using equation (1) we find

$$F = (25 \pm 6) \times 10^3. \tag{2}$$

A proof-of-concept study of chelation in a dry medium

An important step towards the detection of Ba^{2+} in an HPXe is the demonstration that the ions can be chelated in the absence of a solvent. This requires exposing a sample of FBI molecules deposited in a solid–vacuum interface to a source of Ba^{2+} ions.

To achieve this goal, we designed a sublimation experiment as follows. We started by compressing silica powder to form thin silica pellets, then we deposited a FBI–CH$_3$CN solution on the pellets and evaporated the solvent. Two similar SFp pellets (SFpA and SFpB) were prepared by depositing 7.4×10^{-8} mmol of indicator per milligram of silica, which is equivalent to 1.3×10^{15} molecules of FBI. SFpA was kept as a reference for unchelated molecules, and SFpB was introduced in an ultrahigh-vacuum chamber (Fig. 3a) in which barium perchlorate was sublimated. Sublimation was performed using a Knudsen cell at a temperature of around 700 K. The evaporation rate was continuously monitored in situ with a microbalance. The total thickness of deposited $Ba(ClO_4)_2$ was 10 Å, equivalent to a layer of 7.6×10^{14} molecules. Figure 3b, c shows images of the pellet before and after sublimation under an excitation light of 365 nm. The blue shift after sublimation is clearly visible even to the naked eye, showing that a large number of indicators on the pellet's surface were chelated.

The next step was to scan both the SFpA and SFpB pellets in our two-photon absorption (TPA) microscopy setup[44], which is described in some detail in Methods. We performed tomography (for example, Z–X scans) using two filters: a high-pass 'green' filter with $\lambda > 450$ nm and a band-pass 'deep blue' filter with wavelength (400, 425) nm. The Z–X scans were performed with infrared light (800 nm) at a nominal laser power of 100 mW. In addition, we obtained three-dimensional (3D) tomography images, which were assembled from 40 X–Y scans of 75 μm × 75 μm. Each scan corresponded to a different depth Z, in steps of 10 μm. The resulting images were then combined in a 3D image using custom software[45].

Our results are summarized in Fig. 2. We started by measuring the control pellet, SFpA. The Z–X tomography image acquired using the green filter (Fig. 2c) reveals a region of about 20 μm in depth that corresponds to the area of the pellet where FBI molecules were immobilized. Because these are unchelated molecules, they are visible with this filter but not with the deep-blue filter (Fig. 2b). By contrast, for SFpB the green profile (Fig. 2e) is similar to the one measured for SFpA, but the deep-blue tomography (Fig. 2d) shows a clear signal in the same 20-μm region around the pellet surface. This can be exclusively ascribed to the emission of chelated molecules, therefore demonstrating that the sublimation deposited the Ba^{2+} uniformly, resulting in a layer of chelated molecules. Finally, Fig. 2f shows green and deep-blue 3D tomography images confirming that the spatial distribution of the chelated molecules follows that of the unchelated indicators.

Density functional theory (DFT) calculations (described in detail in Methods) show that the Gibbs energy associated with binding of $Ba(ClO_4)_2$ to FBI is −80 kcal mol^{-1}, confirming that the process is very exergonic, which is expected given the experimental result described above and is fully compatible with the high binding constant found for this process.

Chelation of Ba^{2+} by FBI indicators in xenon gas

In an HPXe experiment, the Ba^{2+} created in the $\beta\beta0\nu$ decay will slowly drift to the cathode, picking up on its way neutral xenon atoms in a variety of solvation states, thus yielding $[BaXe_N]^{2+}$ states (with $N = 1, 2...$). At the large pressures that are typical in an HPXe (~20 bar), it has been estimated[46] that $N \approx 8$.

What is the relevance of the proof-of-concept study described here (which demonstrates the observation of the reaction $Ba(ClO_4)_2$ + FBI in vacuo) for an HPXe experiment, which requires that the reaction $[BaXe_N]^{2+}$ + FBI occurs efficiently in high-pressure xenon? DFT can shed light on this question. Our calculations show that the interaction between a Ba^{2+}–Xe_8 cluster and FBI results in a very exergonic process with a calculated Gibbs reaction energy of −195.9 kcal mol^{-1}. This value is almost as large as the Gibbs energy associated with the interaction of a naked dication with the indicator (−197.5 kcal mol^{-1}) and much larger than the energy associated with binding of $Ba(ClO_4)_2$ with FBI (−80.0 kcal mol^{-1}). Furthermore, we find that the Gibbs energy of FBI + Ba^{2+} changes very little in the range 1–30 bar (see Extended Data Table 3).

Finally, our calculations suggest that a layer of indicators with a density of about 1 molecule per square nanometre will efficiently chelate Ba^{2+}. Figure 4 shows the computed structures of FBI and a Ba^{2+}–Xe_8 cluster at different N–Ba^{2+} distances. When optimization of the (**7ca**, Ba^{2+}–Xe_8) pair was started at a N–Ba^{2+} separation of 8 Å, the cluster spontaneously converged to a local minimum at which the original Xe_8 structure was squeezed around the convex face of FBI, and the N–Ba^{2+} distance was 3.27 Å. From this intermediate state, the whole cluster converged to the chelated species, in which the N–Ba^{2+} distance was found to be 2.9 Å. This latter energy minimum was calculated to be about 107 kcal mol^{-1} more stable than the previous intermediate state. In addition, the geometric parameters of the minimum-energy cluster —in which the eight Xe atoms are distributed around FBI—are very similar to those found for the FBI Ba^{2+} and FBI $Ba(ClO_4)_2$ complexes.

Conclusions

We have synthesized an FBI that could be the basis of a barium-tagging sensor in a future HPXe experiment searching for $\beta\beta0\nu$ decays. Using silica as a physical support, we have shown that the FBI has a very large discrimination factor of $F = (25 \pm 6) \times 10^3$ in a dry medium (silica–air). Furthermore, the indicator efficiently chelates Ba^{2+} in a dry medium (silica–vacuum). This was proved by sublimating barium perchlorate ($Ba(ClO_4)_2$) on FBI molecules deposited on a silica pellet and interrogating the indicators using TPA microscopy. To our knowledge, this is the first time that the formation of a Ba^{2+} supramolecular complex in a dry medium is demonstrated.

In addition, we have performed DFT calculations that show that our experimental result is consistent with the exergonic nature of the binding of $Ba(ClO_4)_2$ to the FBI in vacuo and for high solvation states of Ba^{2+} in xenon at all relevant pressures. Importantly, the process evolves spontaneously when the system FBI Ba^{2+} starts at distances of around 1 nm. From these calculations, we can conclude that the formation of supramolecular complexes observed in vacuo implies that FBI indicators can chelate Ba^{2+} ions with high efficiency in an HPXe experiment. We further show in Methods that the large value of F found for the FBI allows the unambiguous identification—using TPA microscopy—of a single chelated indicator.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2431-5.

1. Majorana, E. Theory of the symmetry of electrons and positrons. *Nuovo Cim.* **14**, 171–184 (1937).
2. Gómez Cadenas, J. J., Martín-Albo, J., Mezzetto, M., Monrabal, F. & Sorel, M. The search for neutrinoless double beta decay. *Riv. Nuovo Cim.* **35**, 29–98 (2012).
3. Moe, M. K. New approach to the detection of neutrinoless double beta decay. *Phys. Rev. C* **44**, 931–934 (1991).
4. Danilov, M. et al. Detection of very small neutrino masses in double beta decay using laser tagging. *Phys. Lett. B* **480**, 12–18 (2000).
5. nEXO Collaboration. Imaging individual barium atoms in solid xenon for barium tagging in nEXO. *Nature* **569**, 203–207 (2019).
6. Nygren, D. R. Detecting the barium daughter in ^{136}Xe 0-$\nu\beta\beta$ decay using single-molecule fluorescence imaging techniques. *J. Phys. Conf. Ser.* **650**, 012002 (2015).
7. Jones, B. J. P., McDonald, A. D. & Nygren, D. R. Single molecule fluorescence imaging as a technique for barium tagging in neutrinoless double beta decay. *J. Inst.* **11**, P12011 (2016).
8. McDonald, A. D. et al. Demonstration of single barium ion sensitivity for neutrinoless double beta decay using single molecule fluorescence imaging. *Phys. Rev. Lett.* **120**, 132504 (2018).
9. Sakharov, A. D. Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe. *Pis'ma Z. Eksp. Teor. Fiz.* **5**, 32–35 (1967).
10. Fukugita, M. & Yanagida, T. Baryogenesis without grand unification. *Phys. Lett. B* **174**, 45–47 (1986).
11. Gell-Mann, M., Ramond, P. & Slansky, R. Complex spinors and unified theories. In *Proc. of Supergravity Stony Brook Workshop* (eds Van Nieuwenhuizen, P. & Freedman D. Z.) 27–29 (1979).
12. Yanagida, T. Horizontal symmetry and masses of neutrinos. *Prog. Theor. Phys.* **64**, 1103–1105 (1980).
13. Mohapatra, R. N. & Senjanovic, G. Neutrino mass and spontaneous parity nonconservation. *Phys. Rev. Lett.* **44**, 912–915 (1980).

14. Gando, A. et al. Search for Majorana neutrinos near the inverted mass hierarchy region with KamLAND-Zen. *Phys. Rev. Lett*. **117**, 082503–082506 (2016).
15. GERDA Collaboration. Improved limit on neutrinoless double-β decay of ^{76}Ge from GERDA Phase II. *Phys. Rev. Lett*. **120**, 132503–132505 (2018).
16. Alduino, C. et al. First results from CUORE: a search for lepton number violation via $0\nu\beta\beta$ decay of ^{130}Te. *Phys. Rev. Lett*. **120**, 132501–132508 (2018).
17. Gomez-Cadenas, J. J. Status and prospects of the NEXT experiment for neutrinoless double beta decay searches. Preprint at https://arxiv.org/abs/1906.01743 (2019).
18. Elliott, S. R. & Vogel, P. Double beta decay. *Annu. Rev. Nucl. Part. Sci*. **52**, 115–151 (2002).
19. Sinclair, D. et al. Prospects for barium tagging in gaseous xenon. *J. Phys. Conf. Ser*. **309**, 012005 (2011).
20. Mong, B. et al. Spectroscopy of Ba and Ba$^+$ deposits in solid xenon for barium tagging in nEXO. *Phys. Rev. A* **91**, 022505–022513 (2015).
21. EXO-200 Collaboration. Measurements of the ion fraction and mobility of α- and β-decay products in liquid xenon using the EXO-200 detector. *Phys. Rev. C* **92**, 045504–045510 (2015).
22. Bolotnikov, A. & Ramsey, B. The spectroscopic properties of high-pressure xenon. *Nucl. Instrum. Methods Phys. Res. A* **396**, 360–370 (1997).
23. Nygren, D. High-pressure xenon gas electroluminescent TPC for $0\text{-}\nu\ \beta\beta$-decay search. *Nucl. Instrum. Methods Phys. Res. A* **603**, 337–348 (2009).
24. Álvarez, V. et al. NEXT-100 Technical Design Report (TDR). Executive summary. *J. Inst*. **7**, T06001 (2012).
25. Martín-Albo, J. et al. Sensitivity of NEXT-100 to neutrinoless double beta decay. *J. High Energy Phys*. **2016**, 159 (2016).
26. Thapa, P. et al. Barium chemosensors with dry-phase fluorescence for neutrinoless double beta decay. *Sci. Rep*. **9**, 15097 (2019).
27. Ji, H.-F., Dabestani, R., Brown, G. M. & Hettich, R. Spacer length effect on the photoinduced electron transfer fluorescent probe for alkali metal ions. *Photochem. Photobiol*. **69**, 513–516 (1999).
28. Nakahara, Y., Kida, T., Nakatsuji, Y. & Akashi, M. Fluorometric sensing of alkali metal and alkaline earth metal cations by novel photosensitive monoazacryptand derivatives in aqueous micellar solutions. *Org. Biomol. Chem*. **3**, 1787–1794 (2005).
29. Bissell, R. A. et al. Luminescence and charge transfer. Part 2. Aminomethyl anthracene derivatives as fluorescent pet (photoinduced electron transfer) sensors for protons. *J. Chem. Soc. Perkin Trans*. **2** 9, 1559–1564 (1992).
30. Bourson, J., Pouget, J. & Valeur, B. Ion-responsive fluorescent compounds 4 effect of cation binding on the photophysical properties of a coumarin linked to monoaza- and diaza-crown ethers. *J. Phys. Chem*. **97**, 4552–4557 (1993).
31. Li, J., Yim, D., Jang, W.-D. & Yoon, J. Recent progress in the design and applications of fluorescence probes containing crown ethers. *Chem. Soc. Rev*. **46**, 2437–2458 (2017).
32. Valeur, B. & Berberan-Santos, M. N. in *Molecular Fluorescence – Principles and Applications* 420–436 (Wiley–VCH, 2012).
33. Huston, M. E., Haider, K. W. & Czarnik, A. W. Chelation enhanced fluorescence in 9,10- bis[[(2-(dimethylamino)ethyl)methylamino]methyl]anthracene. *J. Am. Chem. Soc*. **110**, 4460–4462 (1988).
34. Carter, K. P., Young, A. M. & Palmer, A. E. Fluorescent sensors for measuring metal ions in living systems. *Chem. Rev*. **114**, 4564–4601 (2014).
35. Golchini, K. et al. Synthesis and characterization of a new fluorescent probe for measuring potassium. *Am. J. Physiol*. **258**, F438–F443 (1990).
36. Yang, J.-S., Hwang, C.-Y., Hsieh, C.-C. & Chiou, S.-Y. Spectroscopic correlations between supermolecules and molecules. Anatomy of the ion-modulated electronic properties of the nitrogen donor in monoazacrown-derived intrinsic fluoroionophores. *J. Org. Chem*. **69**, 719–726 (2004).
37. Smith, G. A., Hesketh, J. C. & Metcalfe, T. R. Design and properties of a fluorescent indicator of intracellular free sodium concentration. *Biochem. J*. **250**, 227–232 (1988).
38. Crossley, R., Goolamali, Z. & Sammes, P. G. Synthesis and properties of a potential extracellular fluorescent probe for potassium. *J. Chem. Soc. Perkin Trans*. **2** 7, 1615–1623 (1994).
39. Aginagalde, M. et al. Tandem [8 + 2] cycloaddition-[2 + 6 + 2] dehydrogenation reactions involving imidazo[1,2-a]pyridines and imidazo[1,2-a]pyrimidines. *J. Org. Chem*. **75**, 2776–2784 (2010).
40. Zhang, Y., Tang, S., Thapaliya, E. R., Sansalone, L. & Raymo, F. M. Fluorescence activation with switchable oxazines. *Chem. Commun*. **54**, 8799–8809 (2018).
41. Ko, C.-C. & Yam, V. W.-W. Coordination compounds with photochromic ligands: ready tunability and visible light-sensitized photochromism. *Acc. Chem. Res*. **51**, 149–159 (2018).
42. Maitra, R., Chen, J.-H., Hu, C.-H. & Lee, H. M. Synthesis and optical properties of push-push- pull chromophores based on imidazo[5,1,2-cd]indolizines and naphtho[1′,2′:4,5]imidazo[1,2-*a*]pyridines. *Eur. J. Org. Chem*. 5975–5985 (2017).
43. Dougherty, D. A. The cation–π interaction. *Acc. Chem. Res*. **46**, 885–893 (2013).
44. Ávila, F. J., Gambín, A., Artal, P. & Bueno, J. M. In vivo two-photon microscopy of the human eye. *Sci. Rep*. **9**, 10121 (2019).
45. Bueno, J. et al. Multiphoton microscopy of ex vivo corneas after collagen cross-linking. *Invest. Ophthalmol. Vis. Sci*. **52**, 5325–5331 (2011).
46. Bainglass, E., Jones, B. J. P., Foss, F. W., Huda, M. N. & Nygren, D. R. Mobility and clustering of barium ions and dications in high pressure xenon gas. *Phys. Rev. A* **97**, 062509 (2018).

Methods

Photophysics and supramolecular chemistry of FBI indicators in solution

Our experiments to determine the photophysical properties of compounds **7** started by recording their respective emission spectra in acetonitrile solution. Although all compounds were fluorescent with large intensities in the minimum-energy transitions, the critical criterion to select the most suitable candidate was the ability of a given compound to exhibit different lowest emission wavelengths in their unbound and barium-coordinated forms. We defined the peak discrimination factor f_λ at a given wavelength λ as:

$$f_\lambda = \frac{I_\lambda(7\mathrm{Ba}^{2+}) - I_\lambda(7)}{I_\lambda(7)} \tag{3}$$

where $I_\lambda(\mathbf{7}\mathrm{Ba}^{2+})$ and $I_\lambda(\mathbf{7})$ are the intensities of the emission signals at wavelength λ of the corresponding bound ($\mathbf{7}\mathrm{Ba}^{2+}$) and free (**7**) fluorophore. In addition, we measured the molecular brightness[34] B_λ of each transition according to the following expression:

$$B_\lambda = \varepsilon_\lambda \phi_\lambda \tag{4}$$

where ε_λ is the molar extinction coefficient and ϕ_λ is the emission quantum yield.

The data associated with the photophysics of compounds **7** are listed in Extended Data Table 1. According to our results, compound **7aa**, which possesses the 1,4,7-trioxa-10-azacyclododecane moiety (**4a**, $n = 1$), does not show any substantial difference between the free and barium-bound states, thus indicating that this four-heteroatom aza-crown ether is too small to accommodate the Ba^{2+} cation. Compound **7ba**, with a 1,4,7,10-tetraoxa-13-azacyclopentadecane unit (**4b**, $n = 2$), showed a noticeable blue shift upon coordination with Ba^{2+}($\Delta\lambda = -54$ nm). However, the low value of f_λ makes this size of the chelating group not optimal for further development. In the case of the FBI molecule **7ca**, which incorporates the six-heteroatom-containing aza-crown ether unit 1,4,7,10,13-pentaoxa-16-azacyclooctadecane (**4c**, $n = 3$), a larger blue shift associated with Ba^{2+} coordination ($\Delta\lambda = -61$ nm) is observed. Most importantly, the f_λ discrimination factor is found to be of the order of 180, which shows a considerable separation between the unbound **7ca** and the Ba^{2+}-coordinated $\mathbf{7ca}\mathrm{Ba}^{2+}$ species. Both emission spectra are displayed in Extended Data Fig. 1. In addition, both unbound and cationic species show acceptable quantum yields and molecular brightness values.

As far as the chemical structure of the tetracyclic fluorophore is concerned, our results indicate that introducing an additional nitrogen heteroatom in the $2,2\mathrm{a}^1$-diazacyclopenta[*jk*]fluorene to form the corresponding $2,2a^1,3$-triazacyclopenta[*jk*]fluorine analogue is detrimental in terms of quantum yield and molecular brightness, as concluded from the photophysical properties of compound **7da** shown in Extended Data Table 1. Moreover, the presence of an additional fused phenyl group in the fluorophore results in the formation of imidazo[5,1,2-*cd*]naphtho[2,3-*a*]indolizine derivative **7cb**, which has an f_λ factor considerably lower than that measured for **7ca**. Therefore, the presence of additional fused aromatic or heteroaromatic rings to the basic benzo[*a*]imidazo[5,1,2-*cd*]indolizine scaffold does not improve the photophysical properties of the resulting cycloadduct. Finally, the presence of an electron-withdrawing group in compound **7ec** results in a quenching of the quantum yield of the fluorophore, as well as a lowering of the discrimination factor. According to these results, further chemical elaboration of the fluorophore skeleton in order to synthesize the spacer and linker groups shown in Extended Data Fig. 1a must not involve carboxy derivatives such as esters or amides, but π-decoupled moieties such as alkoxy groups. Therefore, we conclude that **7ca** is the optimal combination of structural and electronic features to fulfil our previously defined design criteria.

Having selected compound **7ca** as the best FBI candidate, we conducted studies to assess its binding ability, which must be high (in a dry medium) for our sensor. To that end, we first measured its cation association constant K_a with barium perchlorate in acetonitrile at 298 K using the Benesi–Hildebrand method[47] and the corresponding fluorescence spectra, according to the following formula[48]:

$$\frac{1}{F - F_{\min}} = \frac{1}{F_{\max} - F_{\min}}\left(1 + \frac{1}{K_a[\mathrm{Ba}^{2+}]}\right) \tag{5}$$

In this expression, F is the measured emission of compound **7ca** at the excitation wavelength $\lambda_{\mathrm{exc}} = 250$ nm in the presence of a given $[\mathrm{Ba}^{2+}]$ concentration, and $F_{\min}$ and $F_{\max}$ represent the corresponding intensities of the free aza-crown ether **7ca** and the host–guest complex $\mathbf{7ca}\mathrm{Ba}^{2+}$, respectively. Under these conditions and on the basis of the data shown in Extended Data Fig. 1d, we measured a binding constant of $K_a = 5.26 \times 10^4\ \mathrm{M}^{-1}$ ($R^2 = 0.953$, where R^2 is the coefficient of determination). This indicates the good efficiency of compound **7ca** for Ba^{2+} capture and formation of the $(\mathbf{7ca}\mathrm{Ba}^{2+})(\mathrm{ClO}_4^-)_2$ salt in solution; the favourable photophysical parameters of the compound are listed in Extended Data Table 1. In addition, the Job plot shows a maximum for $n = m = 1$, indicating that **7ca** captures only one Ba^{2+} cation per molecule, as shown in Extended Data Fig. 1e.

Electronic structure calculations and nuclear magnetic resonance experiments

Electronic structure calculations at the DFT level both in the gas phase and in solution confirm the strong binding affinity of **7ca** to coordinate Ba^{2+}. The optimized $\mathbf{7ca}\mathrm{Ba}^{2+}$ structure exhibits a large molecular torsion of the binding group with respect to the free **7ca** molecule (see the dihedral angle ω in Extended Data Fig. 2b) so that a molecular cavity appears, with the metal cation forming a π-complex between the Ba^{2+} metallic centre and the phenyl group. The oxygen atoms of the aza-crown ether occupy five coordination positions with O–Ba contacts within the range of the sum of the van der Waals radii (2.8–3.0 Å)[49]. Interestingly, the phenyl ring attached to the crown ether is oriented towards the centre of the cavity coordinating Ba^{2+} through the π-electrons. The frontier molecular orbitals of **7ca** are delocalized over the entire fluorophore moiety, with virtually no participation of the binding-group electrons (Extended Data Fig. 2c). The lowest bright state of the unbound FBI molecule can be mainly characterized as the electronic transition between the highest occupied molecular orbitals (HOMO) and the lowest unoccupied molecular orbitals (LUMO). Molecular distortion upon metal coordination in $\mathbf{7ca}\mathrm{Ba}^{2+}$ has an important impact on the electronic structure. In particular, the torsion of the phenyl group allowing π-coordination breaks the planarity with the rest of the fluorophore, modifying the HOMO and LUMO energy levels. The decrease of the effective conjugation with respect to **7ca** increases the symmetry allowed $\pi \rightarrow \pi^*$ gap, thus resulting in the blue shift of the fluorescent emission (Extended Data Fig. 2c). Therefore, these results support the viability of **7ca** as an efficient Ba^{2+} indicator in both wet and dry conditions (see Supplementary Information).

Nuclear magnetic resonance (NMR) experiments on the complexation reaction between the FBI molecule **7ca** and barium perchlorate are compatible with the geometries obtained by the DFT calculations. Progressive addition of the salt promoted a deshielding to lower field of the protons of the para-phenylene group marked as *b* in Extended Data Fig. 2d, which are in *ortho* disposition with respect to the aza-crown ether. The *meta* protons marked as *c* in Extended Data Fig. 2d showed a similar, but lower in magnitude, deshielding effect. The remaining protons of the benzo[*a*]imidazo[5,1,2-*cd*]indolizine fluorophore showed a very light deshielding effect but remained essentially unchanged. Instead, the 1,4,7,10,13-pentaoxa-16-azacyclooctadecaane moiety of

7ca showed different deshielding effects upon coordination with Ba^{2+}, with the only exception being the *N*-methylene protons denoted as *a* in Extended Data Fig. 2e, which were shifted to a higher field, thus demonstrating that the nitrogen atom of the aza-crown ether does not participate in the coordination with the dication.

Computed structures of free and complexed FBI

The optimized molecular geometry of the adduct between FBI (**7ca**) and Ba(ClO$_4$)$_2$ (Extended Data Fig. 3) at the DFT level of theory shows a compact structure in which the Ba^{2+} centre does not interact only with the full aza-crown ether but extends its coordination pattern to the N1 atom of the benzo[*a*]imidazo[5,1,2-*cd*]indolizine aromatic tetracycle and to the 1,4-disubstituted phenyl group. Consequently, the nitrogen atom N2 of the aza-crown ether is shifted away from the closest coordination sphere of Ba^{2+} (compare the Ba^{2+}–N1 and Ba^{2+}–N2 distances in Extended Data Table 2). The two perchlorate anions interact with the metallic centre by blocking the extremes of the channel formed by **7ca**, with the Ba^{2+}–O distances only about 0.1 Å larger than those computed for Ba(ClO$_4$)$_2$. This geometry of **7ca**Ba(ClO$_4$)$_2$ results in decoupling between the two components of the fluorophore, with $\omega = 45°$. The calculated Gibbs energy associated with the binding of Ba(ClO$_4$)$_2$ with the FBI is −80 kcal mol^{-1}. This exergonic character is fully compatible with the high binding constant found for this process.

DFT calculations including a naked Ba^{2+} cation bound to **7ca** also showed a rigid structure, in which the main features observed for the **7ca**Ba(ClO$_4$)$_2$ complex—namely, the interaction of the metallic centre with the N1 atom, the oxygen atoms of the aza-crown ethers and the 1,4-disubstituted aromatic ring—are even more pronounced (Extended Data Table 2 and Extended Data Fig. 3). In addition, the reaction is much more exergonic (Gibbs energy of the reaction, $\Delta G_{\mathrm{rxn}} = -197.5$ kcal mol^{-1}; see Extended Data Table 3). The computed energies exhibit a very small dependence on pressure.

If the formation of clusters between the barium cation and the xenon atoms is considered, the interaction of a Ba^{2+}–Xe$_8$ cluster—a species that can be operative under high-pressure conditions—with the FBI results in a still very exergonic process, with a Gibbs reaction energy of −195.9 kcal mol^{-1}. All these results indicate that the findings obtained in solution for the interaction of the FBI compound and barium perchlorate are closely related to the features of the same process in the gas phase involving naked (or Xe-clusterized) barium dications.

Polymer and titration experiments

To measure the response of the FBI in dry media, we studied several materials, including silica (which we selected as our preferred support) and three different polymers: polyvinyl alcohol (PVA), poly(methyl metacrylate) (PMMA) and poly(ether blockamide) (PEBAX 2533).

In the case of silica we conducted a titration experiment, adding increasing concentrations of Ba(ClO$_4$)$_2$ to the gel before depositing the FBI–acetonitrile solution (in each case measurements were performed in a fluorimeter after drying the solvent). Our results are shown in Extended Data Fig. 4a. We found that the response of the complexed FBI indicator improved with larger concentrations of Ba(ClO$_4$)$_2$—an effect that we attribute to the affinity of the silica for barium. For the calculation of *F* we chose the largest concentration studied (7,927 equiv.).

We note, however, that the discrimination factor computed with a concentration of 3,964 equiv. (and with concentrations larger than 7,927 equiv., not shown in the plot) yields a very similar result, compatible with the error quoted for *F*. Our results for the studies with polymers are summarized in Extended Data Fig. 4b, which shows the response of the indicator in PMMA. Under an excitation light of 350 nm, the spectra of both chelated and unchelated molecules are similar and cannot be effectively separated. All the other polymers exhibit a similar behaviour. We attribute the lack of separation between the spectra of chelated and unchelated indicators to the restriction of the conformational freedom imposed by the polymer's rigid environment.

Subtraction of the silica response

Extended Data Figure 5 shows the response of the silica to an excitation light of 250 nm. We note that the subtraction of the silica response results in a zero baseline (and a significant subtraction error) for wavelengths below ~370 nm. Above that value, the chelated spectrum rises quickly, while the unchelated spectrum increases only above ~400 nm. The separation between the two spectra is very large in the region (400, 425) nm, where the response of the uncomplexed spectrum is compatible with zero, but the systematic error in the measurement of the discrimination factor is also large (40%). In the selected region of (400, 425) nm, the separation is still large and the systematic error is reduced to 20%.

Laser setup

A schematic diagram of our laser setup is depicted in Extended Data Fig. 6a. We took advantage of the fact that the emission spectra of the FBI and FBI Ba^{2+} for an excitation light of 250 nm and of 400 nm are very similar (Extended Data Fig. 6b) and used a mode-locked Ti:sapphire infrared laser (800 nm) as the illumination source, inducing the absorption of two photons of 400 nm each. This laser system provided pulses of infrared light with a repetition rate of 76 MHz. The pulse duration was 400 fs on the sample plane. The beam was reflected on a dichroic mirror, passed a non-immersion objective (20×, NA = 0.5) and reached the sample, illuminating a spot limited by diffraction to a volume of about 1 μm^3. A d.c. motor coupled to the objective allowed optical sectioning across the sample along the *Z* direction. This image modality is known as *Z*–*X* tomographic imaging and we call these tomographic images 'profiles'. In addition, we obtained 3D tomography images, which were assembled from 40 *X*–*Y* scans of 75 μm × 75 μm. Each scan corresponded to a different depth *Z*, in steps of 10 μm. The resulting images were then combined in a 3D image. The emitted light was collected through the same objective and passed the dichroic mirror. Finally, before reaching the photomultiplier tube used as the detection unit, the TPA signal passed through either a high-pass, green filter with $\lambda > 450$ nm, or a band-pass deep-blue filter of (400, 425) nm.

To estimate the absolute number of fluorescence photons emitted by the FBI indicator in a TPA scan, we first measured a reference sample of fluorescein suspended in PVA (fluorescein reference sample, FRS). Extended Data Fig. 6c shows a log–log plot of the recorded photomultiplier tube (PMT) signal as a function of the laser power for FRS. As expected for TPA, the slope of the resulting straight line has a value near 2. Extended Data Fig. 6d shows a profile taken on FRS at a power of 80 mW. Identical profiles were taken on SBFp at a power of 40 mW. This allowed the measurement of the brightness ratio $\delta_{\mathrm{r}} = \delta_{\mathrm{SBFp}}/\delta_{\mathrm{FRS}}$, which gave $\delta_{\mathrm{r}} = 17 \pm 4$ and therefore $\delta_{\mathrm{FBIBa}^{2+}} = (6.2 \pm 1.7) \times 10^2$ GM (in units of Goeppert Mayer; 1 GM = 10^{-50} cm^4 s per photon per molecule). The details of the measurement are discussed below.

Determination of the brightness of FBI relative to fluorescein

The fluorophore brightness ($\delta = \sigma\phi_\lambda$, where σ is the TPA cross-section and ϕ_λ is the quantum yield) of fluorescein at a wavelength of 800 nm (ref. [50]) is $\delta_{\mathrm{fluo}} = 36 \pm 9.7$ GM. It is therefore possible to normalize the brightness of the FBI to that of fluorescein by using samples of known concentrations and measuring the response in our setup for identical profiles. To that end, we used a control sample of fluorescein suspended in PVA (FPVA) with a concentration of $n_{\mathrm{FPVA}} = 10^{13}$ molecules cm^{-3} and compared it with an FBI-chelated pellet (SBFp), which had a concentration of $n_{\mathrm{SBFp}} = 2.2 \times 10^{17}$ molecules cm^{-3}. Profiles were taken on FPVA at a power of 500 mW. Identical profiles were taken on SBFp at a power of 100 mW. The total integrated PMT signal in the FPVA and SBFp samples is:

$$I = Kn\delta P^2 \tag{6}$$

where n is the density of molecules (molecules cm^{-3}) of the sample and P is the laser power. K is a constant that depends on the setup, which is the same for the FPVA and SBFp profiles. It follows that:

$$R_{\mathrm{FBI/fluo}} = \frac{\delta_{\mathrm{SBFp}}}{\delta_{\mathrm{FPVA}}} = \frac{I_{\mathrm{SBFp}}}{I_{\mathrm{FPVA}}} \frac{n_{\mathrm{FPVA}}}{I_{\mathrm{SBFp}}} \left(\frac{P_{\mathrm{FPVA}}}{P_{\mathrm{SBFp}}} \right)^2 \tag{7}$$

All the quantities in equation (7) are known. In particular, the integral of the SBFp profile yields 10^9 PMT counts, whereas the integral of the FPVA profile has 5.9×10^4 counts. Thus, we find $R_{\mathrm{FBI/fluo}} = 17 \pm 4$, where the ~20% relative error is dominated by the uncertainty in the concentration n_{SBFp}, and therefore $\delta_{\mathrm{FBIBa^{2+}}} = (6.2 \pm 1.7) \times 10^2$ GM.

Interaction of FBI with other elements

The interaction of the FBI (**7ca**) with other elements was studied in order to assess the selectivity of the indicator. In particular, we chose several dications within the alkaline earth elements, one of which is barium, as well as sodium and potassium, which are abundant in the environment and occupy contiguous positions in the alkaline group of the periodic table. We prepared solutions (5×10^{-5} M) of **7ca** and a metal source in a ratio of 1:1. We used $Ca(OH)_2$, $K(ClO_4)$, $Na(ClO_4)$, $Mg(ClO_4)_2$, $Sr(ClO_4)_2$ and $Ba(ClO_4)_2$ with CH_3CN as the solvent. The results are summarized in Extended Data Fig. 7. We observed that Mg^{2+} induced a partial intensity lowering (on–off effect) at the same emission wavelength upon interaction with **7ca**, whereas Ca^{2+} did not produce any noticeable change in its fluorescence emission spectrum when mixed with **7ca**. Therefore, we concluded that our indicator does not produce substantial changes to the emission wavelength in the presence of light alkaline earth dications. By contrast, in the presence of **7ca**, Sr^{2+} exhibited an emission spectrum similar to that observed for Ba^{2+}. These results show that **7ca** is able to chelate the heavier alkaline earth dications Sr^{2+} and Ba^{2+}. It is therefore expected that **7ca** should chelate Ra^{2+}. Finally, according to our results, neither K^+ nor Na^+ were chelated by **7ca**, thus evidencing the high selectivity of our indicator.

A sensor for Ba^{2+} tagging

In addition to a sensor capable of chelating Ba^{2+} with high efficiency, a future HPXe experiment with barium tagging needs to be able to distinguish unambiguously the signal of a single complexed indicator from the background of unchelated surrounding molecules. Here we show that the large discrimination factor of the FBI permits such a robust observation of single chelated molecules even for densely packed sensors.

We consider a TPA microscopy system similar to the one used here, but with optimized parameters, for example, an 800-nm pulsed laser, with a repetition rate of $f = 100$ MHz, pulse width $\tau = 100$ fs full-width at half-maximum and a moderately large numerical aperture of NA = 0.95. Following ref. [51], we take the overall light collection efficiency of the system to be $\varepsilon_c = 10\%$. Focusing the laser on a diffraction-limited spot (a circle of ~0.5 μm diameter) results in a photon density of 1.7×10^{31} photons cm^{-2} W^{-2} per pulse.

We assume now that a single FBI molecule complexed with a Ba^{2+} ion and m unchelated indicators are contained in such a diffraction-limited spot. The number of absorbed photons, n_a, per fluorophore and per pulse is[52]:

$$n_a = \frac{P^2 \delta}{\tau f} \left(\frac{\mathrm{NA}^2}{2 \hbar c \lambda} \right)^2 \tag{8}$$

where P is the laser power, δ is the brightness ($\sigma\phi_\lambda$) of the fluorophore, $\hbar$ is the reduced Planck constant and c is the speed of light in vacuum.

We can compute the number of photons that the chelated indicator absorbs as a function of the laser power using equation (8). Given the relatively large TPA cross-section of the FBI (also computed here), $n_a = 2$ for a modest power of 11 mW. By setting the laser power at this value, the emission rate of the chelated molecule will equal the laser repetition rate, $n_f = 1 \times 10^8$ photons s^{-1}.

The light emitted by the complexed FBI molecule will be blue-shifted. We assume that a band filter λ_f of (400, 425) nm is placed in front of the CCD. n_f is the fluorescence emitted in a given time interval by the chelated indicator. Then, the light recorded by the CCD that is due to the chelated indicator will be $N = \varepsilon_f \varepsilon_c n_f$, where $\varepsilon_f \approx 0.29$ is the band-pass filter efficiency for the signal.

The total fluorescence (green-shifted) emitted by the unchelated molecules will be $m n_f / C$, and the corresponding background light recorded by the CCD will be $N_b = \varepsilon_f' \varepsilon_c m n_f / C$, where $\varepsilon_f' \approx 0.0036$ is the band-pass filter efficiency for the background.

The total signal N_t recorded in the CCD will be $N_t = N_f + N_b$, where N_f is the fluorescent signal. The estimator of the signal observed in the spot will be $N_t - N_b$, where N_b can be computed with great precision by taking the average of a large number of spots containing only unchelated molecules. The signal-to-noise ratio (SNR) of the subtraction is:

$$\mathrm{SNR} = \frac{N_f}{\sqrt{N_b}} = \sqrt{\varepsilon_f \varepsilon_c \frac{n_f F}{m}} = \sqrt{\frac{7.2 \times 10^{10}}{m}} \tag{9}$$

in units of s$^{1/2}$. The SNR is expressed as a function of time in seconds because n_f measures the number of photons per second. The number of molecules in the diffraction spot will depend on the density of indicators, ρ, in the sensor. We assume that the target will be a dense monolayer with about one molecule per square nanometre. As shown by our DFT calculations, the 'snowballs' formed by the barium ion during transport (for example, $Ba^{2+}Xe_8$) will readily form a supra-molecular complex at distances of the order of 1 nm (for example, 8 Å in the example discussed here). Thus, $\rho = 10^6$ μm^{-2} and $m = 2 \times 10^5$. By substituting in equation (9), we find SNR = 6×10^2 s$^{1/2}$. If we take a scanning time per spot of 1 ms, then SNR ≈ 20. Therefore, a chelated indicator would produce an unmistakable signal above the background of unchelated molecules in that spot. This demonstrates that fast and unambiguous identification of Ba^{2+} ions in the sensor can be attained using a dense monolayer. The scanning of large surfaces using wide-field TPA is discussed below.

A BOLD concept

We conceive the Barium atOm Light Detector (BOLD), which is an HPXe implementing a full barium-tagging detector (BTD) that fully covers the cathode of the apparatus. Other possibilities that could apply to a future HPXe with barium tagging are discussed in ref. [53].

BOLD consists of three major systems. An energy-tracking detector (ETD), which measures the energy and the start time t_0 of the event and reconstructs its topology (and in particular its barycentre), and the BTD, which is capable of tagging, with high efficiency, the single Ba^{2+} ion produced in a $\beta\beta0\nu$ or $\beta\beta2\nu$ decay. The information of these two systems is linked through the delayed coincidence trigger (DCT), which establishes a coincidence between the observation of the two-electron signal and the detection of Ba^{2+}. The role of the DCT is to suppress the impact of $\beta\beta2\nu$ events and of other potential accidental coincidences involving ions such as Ra^{2+} and Sr^{2+}.

Extended Data Fig. 8 shows a schematic of BOLD. Conceptually, the detector is as follows: the ETD is an array of light sensors (probably silicon PMTs) located behind the transparent anode, which is connected to high voltage. The BTD is located behind the grounded cathode and deploys an array of tiles called Molecular Target Elements (MTEs). A self-assembled monolayer of FBI indicators is grown on one of the sides of the MTEs, and placed facing the TPC fiducial volume. The MTEs are interrogated by a fast TPA laser microscopy system (TPAL) consisting of one or more pressure-resistant objectives, which are able to move on demand to the specific area of the BTD that needs to be scanned. The laser will be a high-power (2–3 MW), pulsed, femtosecond, 100-MHz (or 1-GHz) system that enters the chamber through suitable windows and is steered by piezo-electric actuated mirrors.

A prototype of such a system is already under development as a part of the NEXT R&D programme[53].

The delayed coincidence trigger is activated by the ETD when the energy of the event is measured to be within the region of interest, signalling an event of interest. When this happens, the ETD reconstructs the barycentre of the event and computes the expected time of arrival of the Ba^{2+} ion to the BTD. It then sends the coincidence trigger, which lowers the voltage of the BTD during a time window large enough (about 1 ms) to allow the putative Ba^{2+} ion arriving to the cathode to 'cross the gate', reach the BTD and be captured by one of the MTEs. The predicted arrival position of the ion is also known from the barycentre of the event (with a resolution of about 5 mm at a pressure of 40 bar, according to our Monte Carlo calculations) and is sent to the TPAL, which scans a region around it. After scanning, the TPAL sends a signal if a chelated molecule has been found. The signature of a $\beta\beta0\nu$ event is the coincidence between the energy trigger, the time trigger opening the cathode gate, and the TPAL positive trigger.

Given the barycentre resolution of 5 mm, the Ba^{2+} candidate will be contained in a scanning region of 1.5 cm × 1.5 cm more than 99% of the times. To scan such an area in a reasonable time, it is necessary to implement large-field-of-view (FOV) techniques. For example, a FOV of 100 μm diameter and an interrogation rate of 1 ms per FOV result in a scanning time of 13 s cm^{-2}, which allows the scanning of the barium fiducial area (1.5 cm × 1.5 cm) in ~30 s.

Indeed, the availability of lasers with peak powers of several watts makes fast scanning possible by using wide-field two-photon microscopy[54]. If, instead of focusing into a diffraction-limited spot by overfilling the back aperture of the objective (as discussed in the example given in the main text), we choose to focus into a small spot near the back aperture, a wider (and weaker) spot is produced on the target plane. The number of absorbed photons in this configuration decreases with $(r/r_d)^2$, where r is the wide-field radius and r_d is the radius of the diffraction limit spot. By taking $r = 50$ μm and $r_d = 0.5$ μm, we find that $n_a^{wf} = n_a \times 10^{-4}$, where n_a^{wf} is the number of absorbed photons in the wide-field configuration. However, these four orders of magnitude can be accounted for by the P^2 dependence of n_a. Indeed, we find that $n_a^{wf} = 2$ for a power of 2.1 W. By projecting each diffraction-limited spot in the FOV in one CCD pixel, it is then possible to find whether any pixel in the CCD has a chelated molecule with high SNR (~20) in 1 ms (the last generation of CCD cameras features speeds in excess of 1,000 frames per second), and thus fast scanning is feasible.

The scanning methodology deserves also some comments. During the fabrication of the BTD, each of the MTEs will be scanned and a map of pixels will be recorded. The map will contain the position of the pixel and the intensity response in the deep-blue band (for example, a filter of (400, 425) nm) to the interrogation of the scanning laser operating at the nominal parameters. The initial scan will allow us to identify and reject defective MTEs and to veto any potential defective spots.

Under normal operation, when the DCT triggers the scan of a specific region, the system records the signal in each spot and compares it with the reference in the database, as well as with the running average computed in real time. This allows us to take into account any local variation of density in the MTEs, as well as fluctuations in the laser power, which are controlled with very good precision. The systematic error that we obtain when simulating these parameters is small compared with the bulk effect of the subtraction of the light that is due to unchelated molecules. Setting a very high nominal SNR (20 in our analysis) also provides extra protection against spurious fluctuations, which in our analysis never yielded an SNR greater than 3. To conclude, we think that a robust and reliable TPA scanning system can be implemented.

Data availability

The data that support the findings of this study are available within the paper and Supplementary Information. Additional data generated during the present study are available from the corresponding authors upon reasonable request.

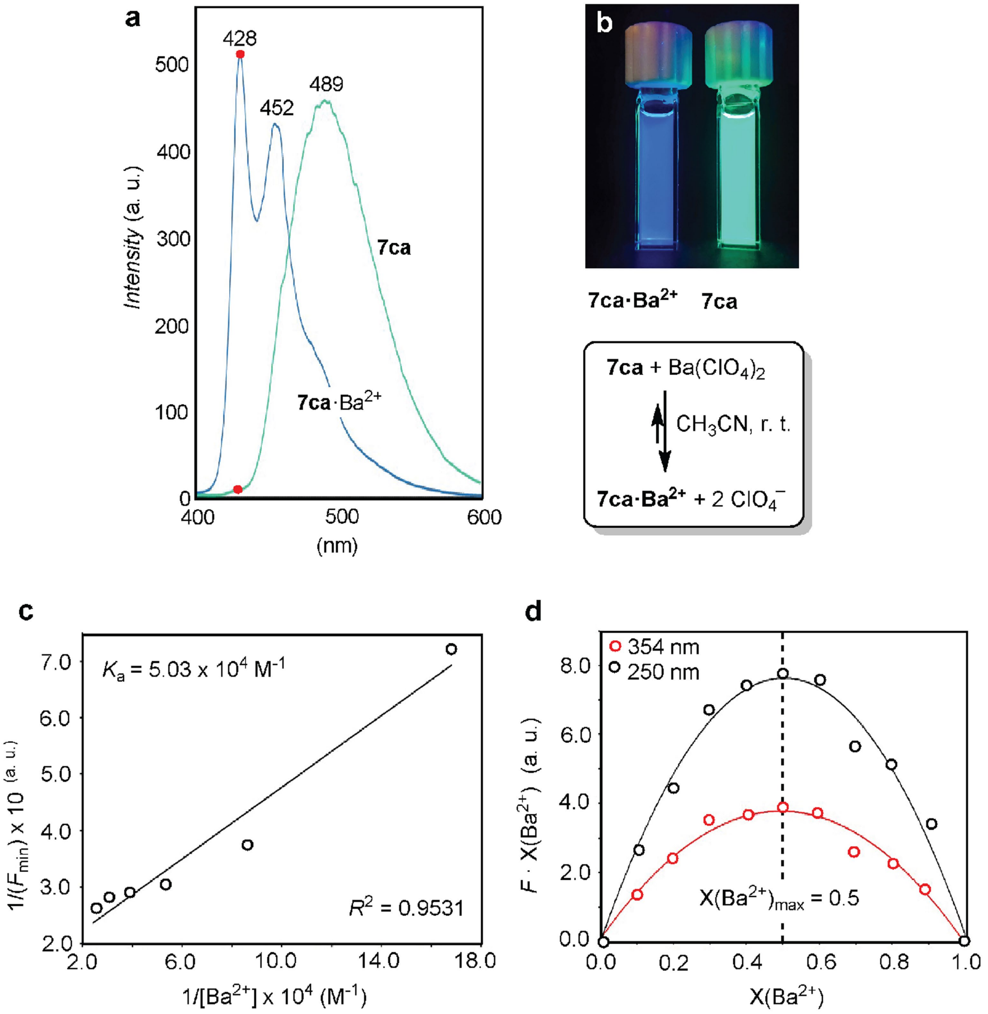
47. Benesi, H. A. & Hildebrand, J. H. A spectrophotometric investigation of the interaction of iodine with aromatic hydrocarbons. *J. Am. Chem. Soc.* **71**, 2703–2707 (1949).
48. Zhang, Q. & Duan, K. Fluorescence chemosensor containing 4-methyl-7-coumarinyloxy, acetylhydrazono and n-phenylaza-15-crown-5 moieties for K^+ and Ba^{2+} ions. *Heterocycl. Commun.* **24**, 141–145 (2018).
49. Batsanov, S. S. Van der Waals radii of elements. *Inorg. Mater.* **37**, 871–885 (2001).
50. Xu, C. & Webb, W. W. Measurement of two-photon excitation cross sections of molecular fluorophores with data from 690 to 1050 nm. *J. Opt. Soc. Am. B* **13**, 481–491 (1996).
51. Zinter, J. P. & Levene, M. J. Maximizing fluorescence collection efficiency in multiphoton microscopy. *Opt. Express* **19**, 15348–15362 (2011).
52. Denk, W., Strickler, J. H. & Webb, W. W. Two-photon laser scanning fluorescence microscopy. *Science* **248**, 73–76 (1990).
53. Byrnes, N. K. et al. Barium tagging with selective, dry-functional, single molecule sensitive on-off fluorophores for the NEXT experiment. In *Meeting of the Division of Particles and Fields of the American Physical Society* https://www.slac.stanford.edu/econf/C1907293/ (2019).
54. Amor, R. et al. Widefield two-photon excitation without scanning: Live cell microscopy with high time resolution and low photo-bleaching. *PLoS One* **11**, e0147115 (2016).

Acknowledgements We acknowledge the support of our colleagues in the NEXT collaboration in the development of this work as a part of the R&D programme to develop a background-free experiment based on Ba^{2+} tagging. We also acknowledge support from the following agencies and institutions: the European Research Council (ERC) under Advanced Grant 339787-NEXT; the Ministry of Science and Innovation of Spain and FEDER under grants FIS2014-53371-C04, FIS2016-76163-R, MAT2016-78293-C6-5-R, MINECO/FEDER CT2016-80955-P, CTQ2016-80375-P and CTQ2014-51912-REDC; Interred PCTEFA V-A Spain/France/Andorra Program (EFA 194/16/TNSI); the Basque Government (GV/EJ) under grants IT-1346-19 and IT-1180-19; and Agencia de Ciencia y Tecnología de la Región de Murcia (19897/GERM/15). The authors also thank the SGI/IZO-SGIker UPV/EHU, Fundación Séneca and DIPC for computational and analytical resources.

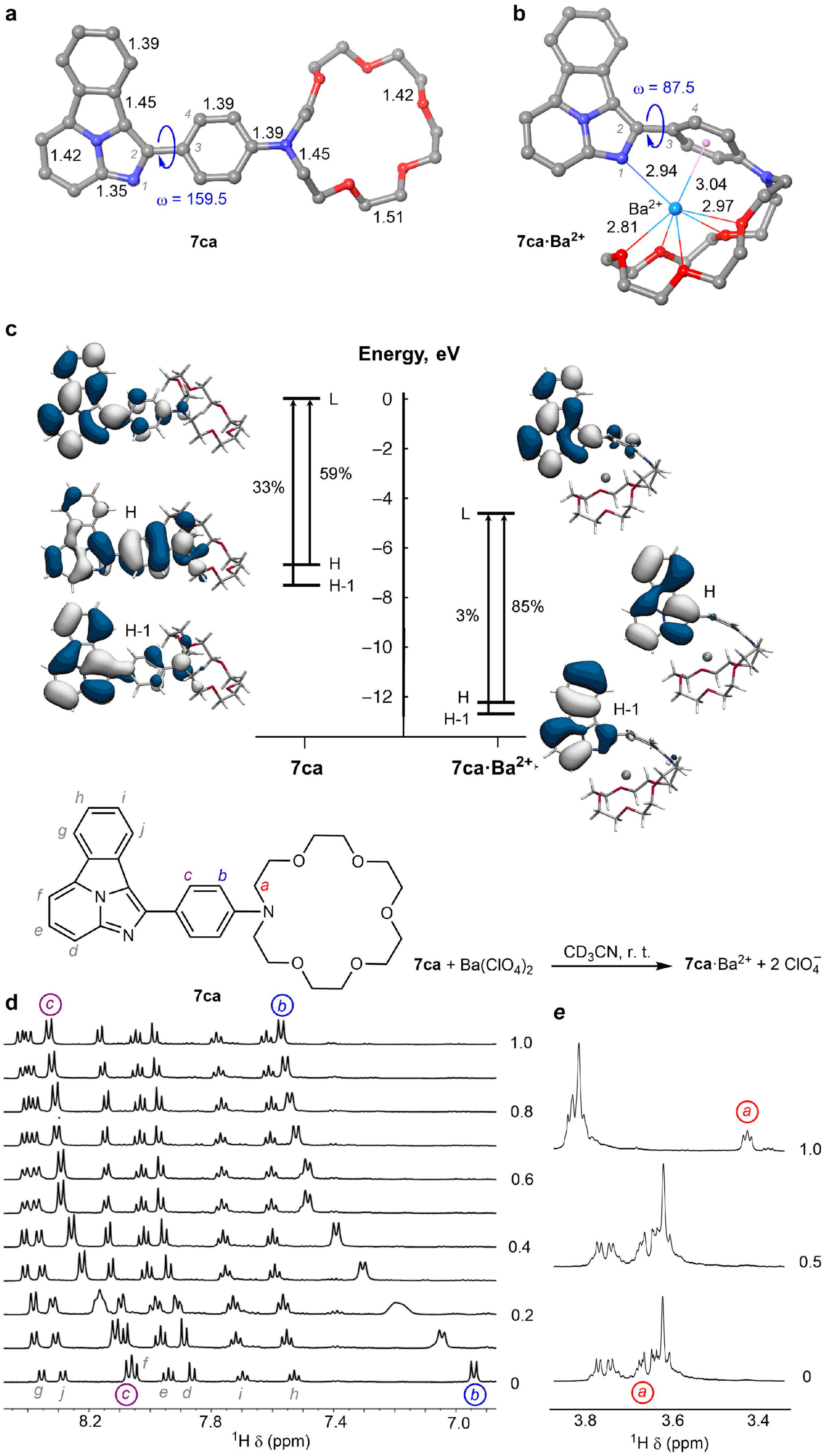
Author contributions J.J.G.-C., F.P.C. and D.N. conceived the project. J.J.G.-C. and F.P.C. coordinated the experiments and analysed the data. I.R. and B.A. carried out the chemical synthesis, characterization and solution fluorescence studies of the compounds. J.I.M. carried out the NMR experiments. C.T., F.P.C. and D.C. performed the computational studies. P.H. and C.R. designed and performed the chelation in dry medium. Z.F. carried out the fluorescence studies in the silica experiments. B.O. and T.S. performed the solid-phase experiments involving polymers. J.M.B., R.M.M.-O., P.H., F.M. and P.A. performed the laser experiments (coordinated by J.M.B.). J.J.G.-C. and F.P.C. wrote the manuscript. D.N., F.M., C.R. and Z.F. assisted in writing and editing the manuscript.

Competing interests The authors declare no competing interests.

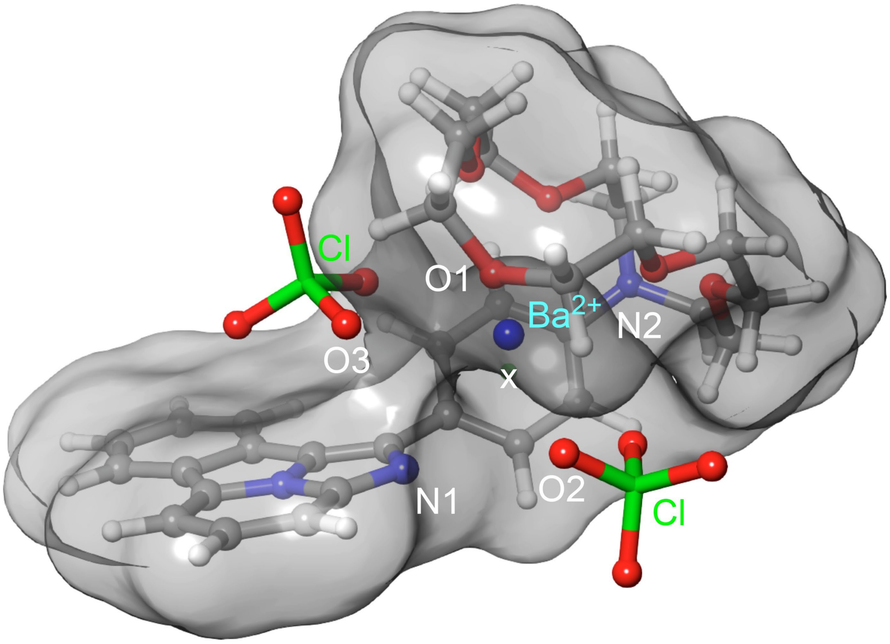
Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2431-5.
Correspondence and requests for materials should be addressed to F.P.C. or J.J.G.-C.
Peer review information *Nature* thanks Mark Chen and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.
Reprints and permissions information is available at http://www.nature.com/reprints.



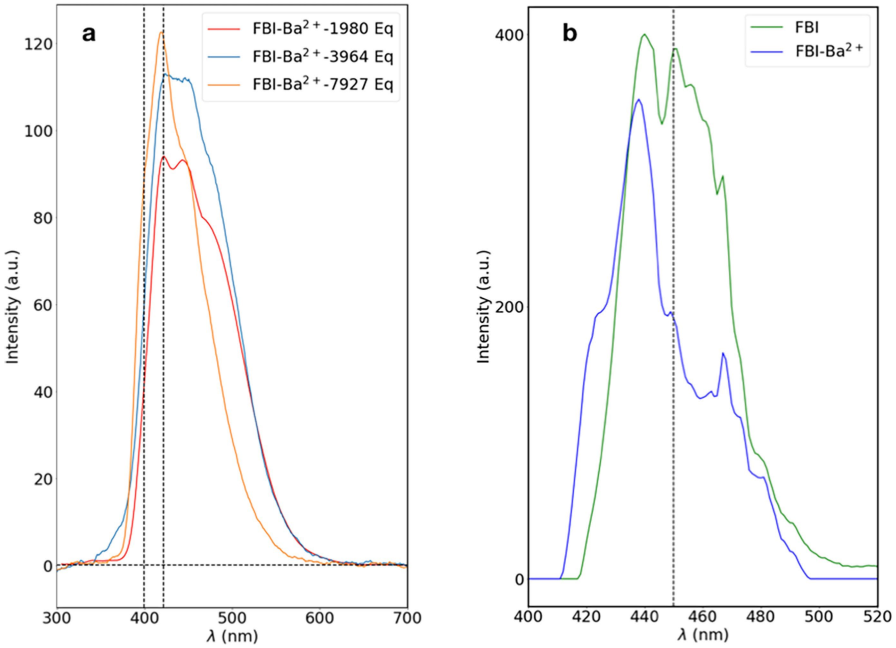
Extended Data Fig. 1 | Characterization of FBI in solution. **a**, Emission spectra of unchelated (**7ca**; cyan) and chelated (**7ca**Ba^{2+}; blue) indicators upon excitation at 250 nm. Red dots indicate the wavelengths used to determine the peak discrimination factor f_λ. **b**, Photographs of the two species in acetonitrile showing bicolour emission upon irradiation at 365 nm. **c**, Benesi–Hildebrand plot of the fluorescence emission spectra of FBI in acetonitrile solution at room temperature in the presence of different concentrations of barium perchlorate. **e**, Job's plot of the **7ca** + Ba(ClO$_4$)$_2$ interaction, showing a 1:1 stoichiometry between **7ca** and Ba^{2+}, thus forming the complex **7ca**Ba^{2+}. ΔF, variation in the measured emission; X(Ba^{2+}), molar fraction of Ba^{2+}.



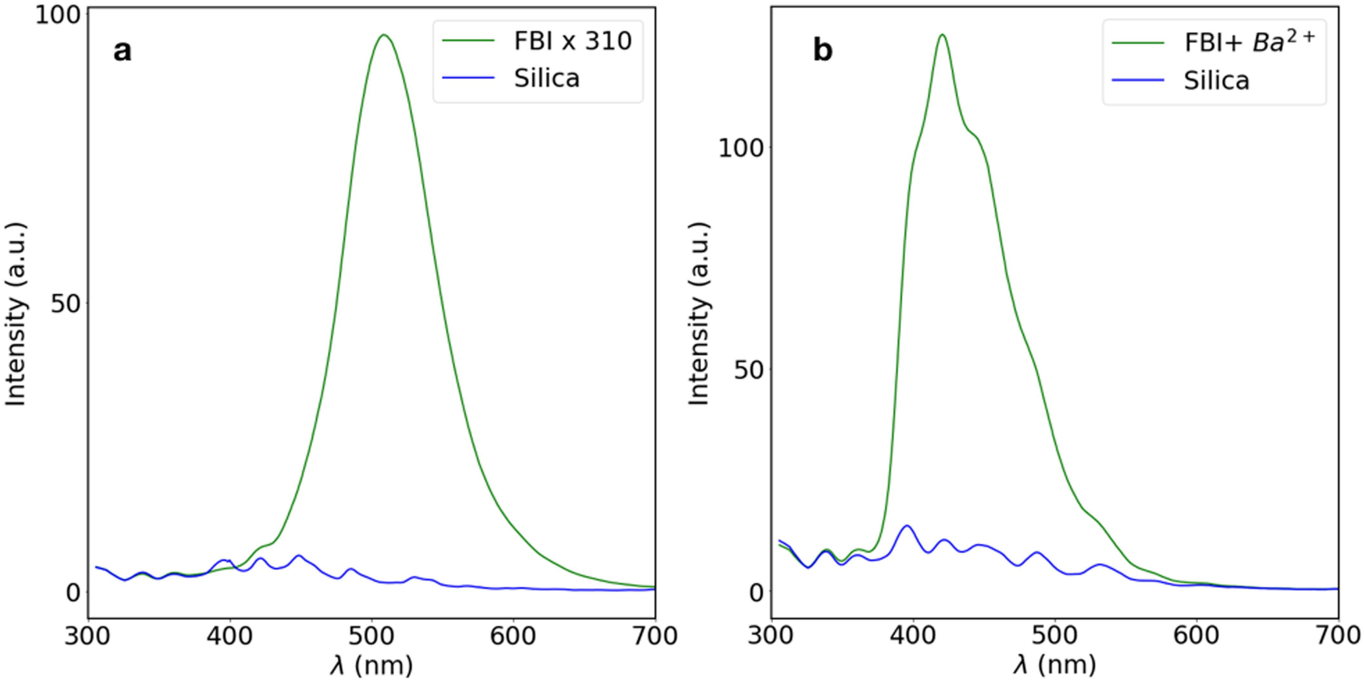
Extended Data Fig. 2 | Theoretical predictions and NMR experiments. **a**, **b**, DFT-derived gas-phase structures of **7ca** (**a**) and **7ca**Ba^{2+} (**b**). Bond distances are given in Å. Dihedral angles *ω* formed by covalently bonded atoms 1–4 are given in degrees and in absolute values. **c**, Frontier molecular orbital energy diagram of **7ca** (left) and **7ca**Ba^{2+} (right). Vertical arrows indicate the main contributions to the electronic transition to the lowest bright state. **d**, **e**, Aromatic (**d**) and aza-crown ether (**e**) regions of the proton NMR spectra of compound **7ca** upon addition of barium perchlorate. The most important changes in chemical shift (in ppm) are highlighted. All the spectra were recorded at 500 MHz. Protons *a* correspond to the methylene groups of the aza-crown ether moiety (**e**). Protons *b* and *c* (**d**) correspond to the para-benzylidene group. See the drawing of **7ca** in **d** for the assignment of all protons.



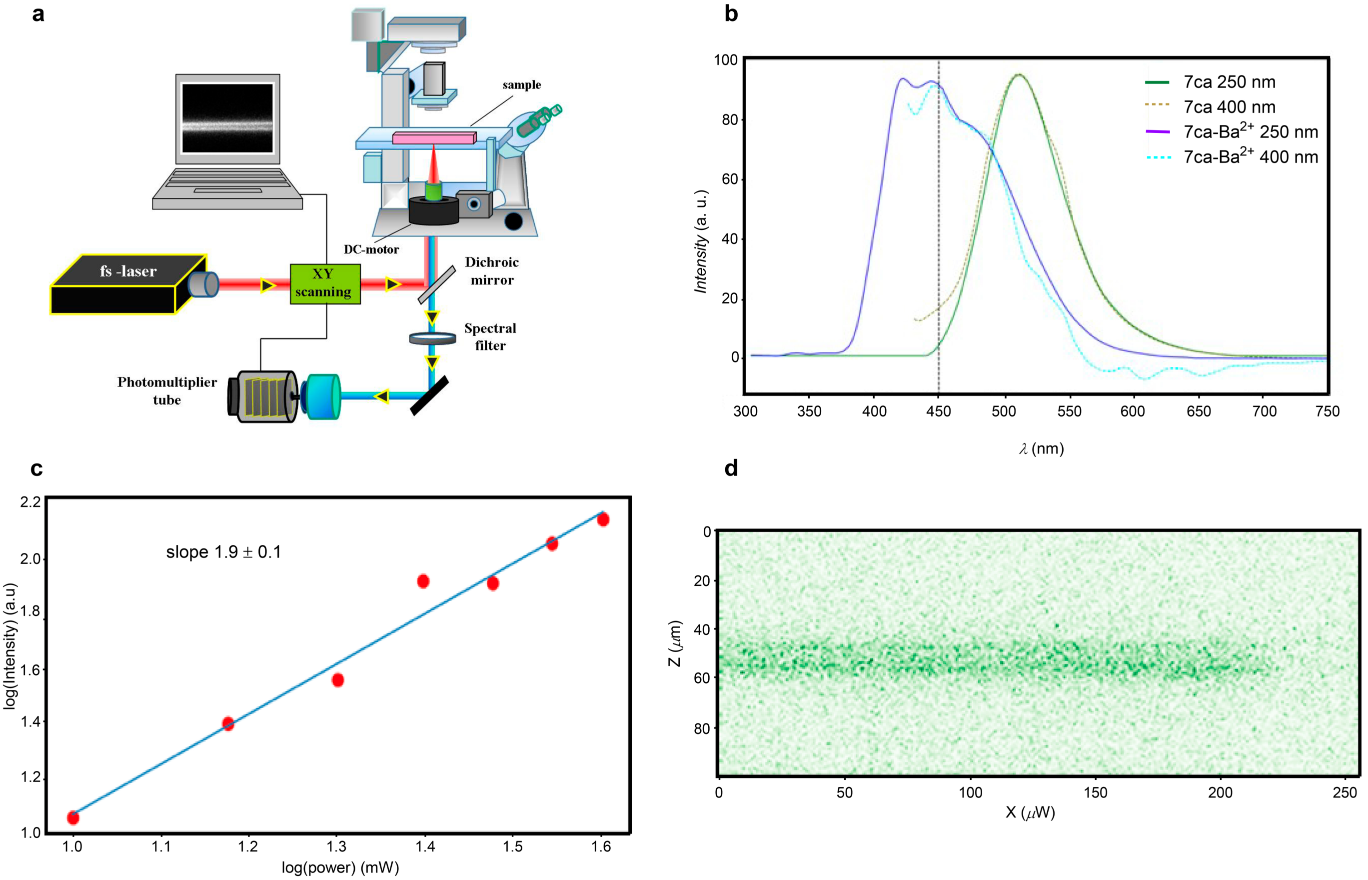
Extended Data Fig. 3 | Computed structures of FBI–barium perchlorate complex. DFT-derived fully optimized structure of **7ca** complexed with barium perchlorate. A dummy atom located at the centre of the 1,4-disubstituted phenyl group is denoted as X. Bond distances and dihedral angles are given in Extended Data Table 2.



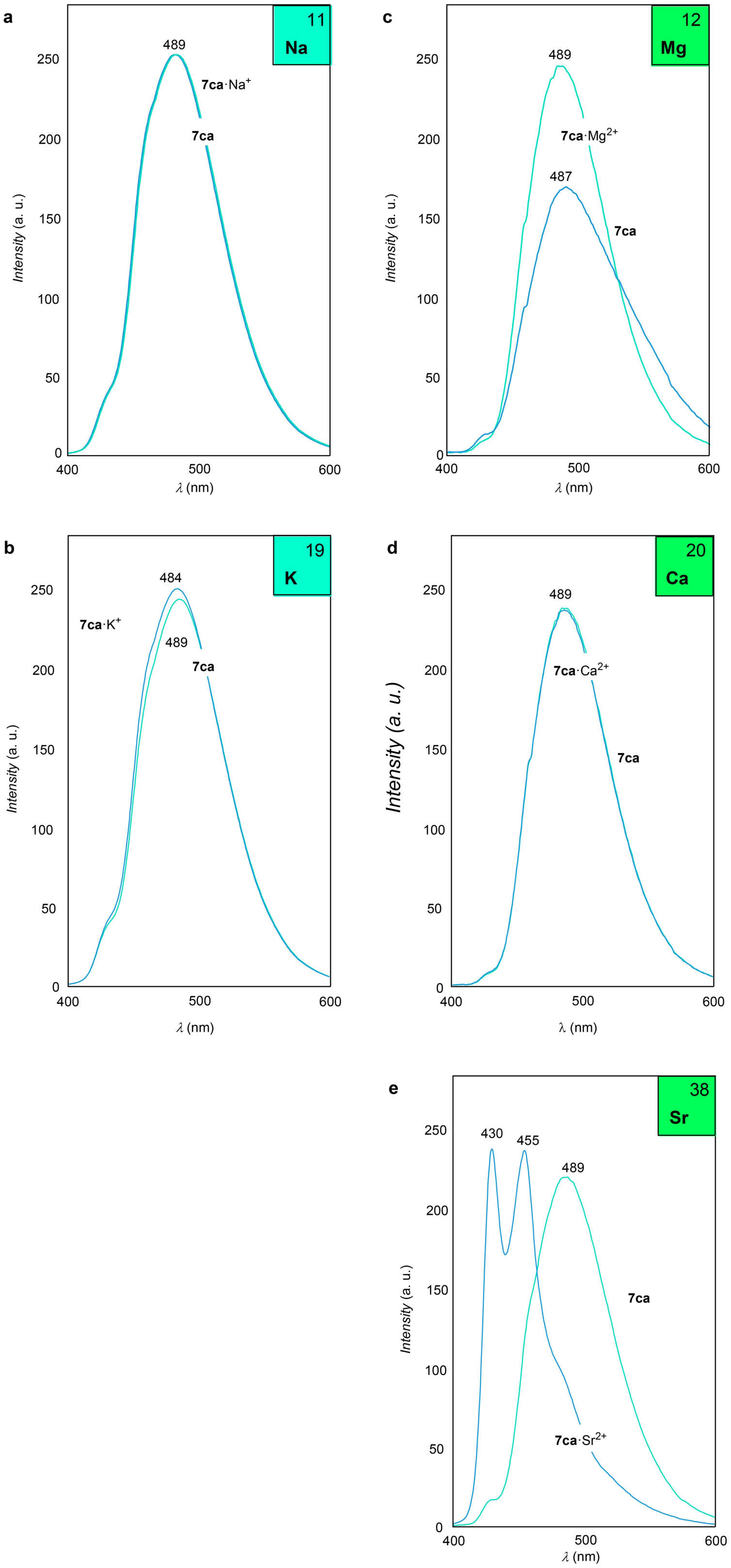
Extended Data Fig. 4 | Titration and polymer experiments. a, Titration experiments, showing that the response of the FBI improves for larger concentrations of barium. Eq, equivalent. **b**, Example of a polymer experiment, showing that the response of the FBI loses its characteristic colour shift.



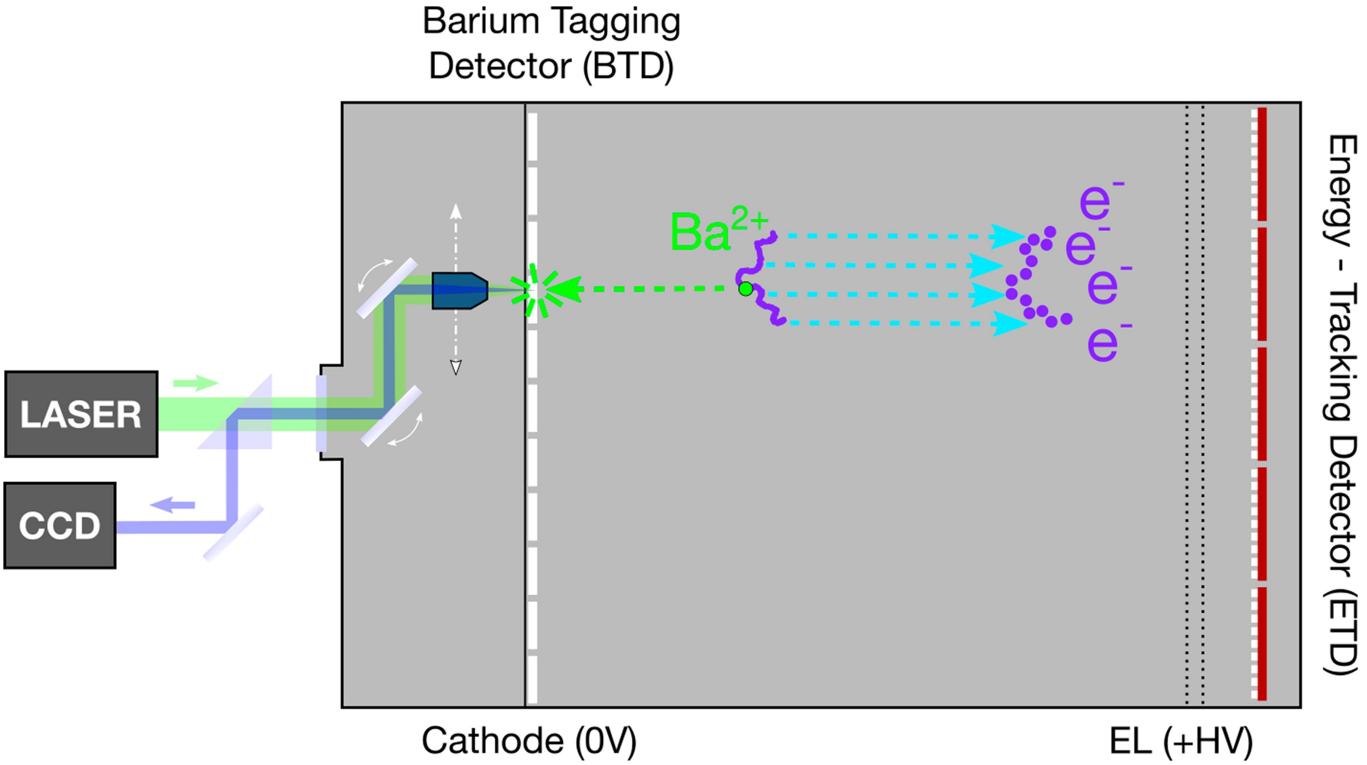
Extended Data Fig. 5 | Subtraction of the silica response. a, **b**, Emission spectra of the SF (**a**) and SBF (**b**) samples, with the background from the silica superimposed, for an excitation light of 250 nm.



Extended Data Fig. 6 | TPA microscopy. a, Illustration of our setup. An infrared (800 nm) laser passes through a dichroic mirror and fills the back plane of the objective (20×, NA = 0.5) of an inverted microscope. The laser is focused in the sample, with a spot limited by diffraction (for example, a volume of about 1 μm^3). The emitted fluorescence passes through a selection filter before being recorded by a PMT. **b**, Emission spectra of FBI and FBI Ba^{2+} for an excitation light of 250 nm (green, blue) and 400 nm (olive, cyan). The spectra are very similar, allowing the use of an infrared laser of 800 nm for our proof-of-concept study. **c**, Log–log plot showing the quadratic dependence of the intensity on the power, which is characteristic of TPA, for the FRS. **d**, Two-dimensional scan (profile) across the FRS. Integration of the profile yields an integrated signal that can be used for the normalization of the FBI samples.



Extended Data Fig. 7 | Interaction of FBI with other elements (1:1 equiv.). **a**–**e**, Blue lines represent FIB + Na$^+$ (**a**), FIB + K$^+$ (**b**), FIB + Mg^{2+} (**c**), FIB + Ca^{2+} (**d**) and FIB + Sr^{2+} (**e**), and the cyan lines show the corresponding unchelated indicators. In **a**–**d**, the spectra show that the FIB is not chelated with the ion, whereas in **e** the response is similar to that observed for barium, showing the formation of a supramolecular complex. All excitation spectra were taken at 250 nm.



Extended Data Fig. 8 | Schematic of the BOLD detector. An example of a $\beta\beta0\nu$ signal event is shown. The two electrons emitted in the decay (purple) propagate in the dense xenon gas ionizing it, and the ionization electrons drift towards the anode, where their energy is measured by the ETD, which also reconstructs the event barycentre. The Ba^{2+} ion drifts very slowly towards the cathode, where it is eventually captured and identified by the BTD.

Extended Data Table 1 | Characterization of FBI compounds **7** and **7**Ba^{2+}

Compound	λ_{em} ^{a} (nm)		f_λ ^{b}	ϕ_λ ^{c}		B_λ ^{d} (M^{-1}cm^{-1})	
	7	7—Ba^{2+}	7—Ba^{2+}	7	7—Ba^{2+}	7	7—Ba^{2+}
7aa	485	485	0.07	0.42	0.41	8.42	8.45
7ba	482	428	6.02	0.34	0.32	7.65	8.13
7ca	489	428	179.74	0.67	0.45	11.26	8.06
7da	491	491	n. d.	0.06	0.06	0.53	0.51
7ec	511	430	22.64	0.29	0.25	3.65	3.05
7cb	503	456	4.86	0.22	0.04	4.84	1.21

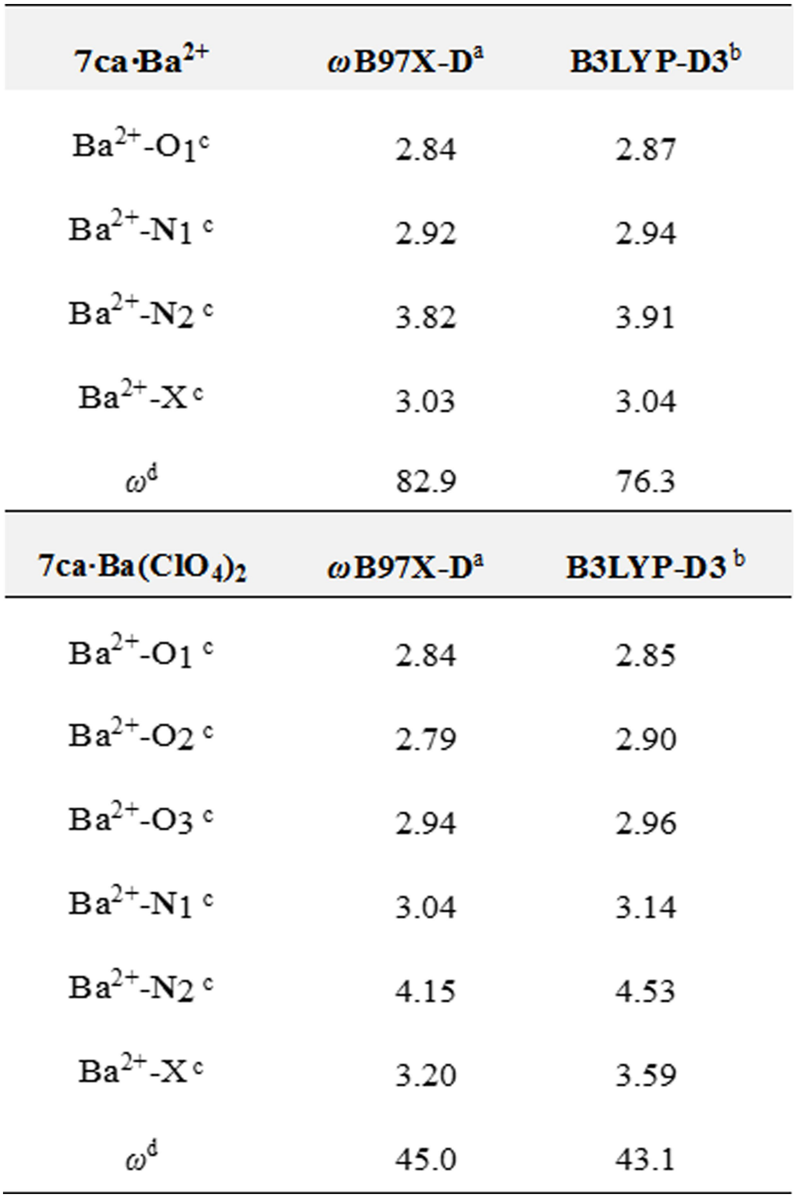
[a] Emission wavelengths (λ_{em}) at an excitation wavelength of 250 nm.

[b] Peak discrimination factors (f_λ) with respect to unbound fluorophores **7** at λ_{em}. n. d., not determined.

[c] Quantum yields (ϕ_λ) at λ_{em}.

[d] Molecular brightness of the fluorescent emissions (B_λ) at λ_{em}.

Extended Data Table 2 | Structural parameters for the geometries of 7caBa^{2+} and 7caBa(ClO$_4$)$_2$



7ca·Ba^{2+}	**ωB97X-D**[a]	**B3LYP-D3**[b]
Ba^{2+}-O1[c]	2.84	2.87
Ba^{2+}-N1[c]	2.92	2.94
Ba^{2+}-N2[c]	3.82	3.91
Ba^{2+}-X[c]	3.03	3.04
ω[d]	82.9	76.3

7ca·Ba(ClO$_4$)$_2$	**ωB97X-D**[a]	**B3LYP-D3**[b]
Ba^{2+}-O1[c]	2.84	2.85
Ba^{2+}-O2[c]	2.79	2.90
Ba^{2+}-O3[c]	2.94	2.96
Ba^{2+}-N1[c]	3.04	3.14
Ba^{2+}-N2[c]	4.15	4.53
Ba^{2+}-X[c]	3.20	3.59
ω[d]	45.0	43.1

The atomic labels are shown in Extended Data Fig. 3.

[a] Structures optimized in vacuo using DFT.

[b] Structures optimized in vacuo using DFT.

[c] Bond distances in Å.

[d] Dihedral angles (absolute value) are given in degrees.

Extended Data Table 3 | Gibbs reaction energies of compound 7ca with Ba^{2+} under different conditions

Reaction	Pressure (atm)	ΔG_{rxn}[a] (kcal/mol)
7ca + Ba^{2+} → **7ca**−Ba^{2+}	1	-197.5
	10	-198.8
	20	-199.2
	30	-199.5
7ca + Ba^{2+}−8Xe → **7ca**−Ba^{2+}+8Xe	1	-195.9[b]
7ca + $Ba(ClO_4)_2$ → **7ca**−$Ba(ClO_4)_2$	1	-80.0

[a] ΔG_{rxn} is the free energy of the reaction, calculated as $\Delta G_{rxn} = \Sigma G_{prod} - \Sigma G_{react}$ (ΣG_{prod}, total free energy of the product; ΣG_{react}, total free energy of the reactants) at a temperature of 298.15 K and computed using DFT.

[b] Free energy of the reaction, computed considering isolated **7ca**Ba^{2+} clusters and eight individual Xe atoms as reaction products.

Laser picoscopy of valence electrons in solids

https://doi.org/10.1038/s41586-020-2429-z

Received: 29 April 2019

Accepted: 1 May 2020

H. Lakhotia[1,2,4], H. Y. Kim[1,2,4], M. Zhan[1,2], S. Hu[3], S. Meng[3] & E. Goulielmakis[1,2]✉

Valence electrons contribute a small fraction of the total electron density of materials, but they determine their essential chemical, electronic and optical properties. Strong laser fields can probe electrons in valence orbitals[1–3] and their dynamics[4–6] in the gas phase. Previous laser studies of solids have associated high-harmonic emission[7–12] with the spatial arrangement of atoms in the crystal lattice[13,14] and have used terahertz fields to probe interatomic potential forces[15]. Yet the direct, picometre-scale imaging of valence electrons in solids has remained challenging. Here we show that intense optical fields interacting with crystalline solids could enable the imaging of valence electrons at the picometre scale. An intense laser field with a strength that is comparable to the fields keeping the valence electrons bound in crystals can induce quasi-free electron motion. The harmonics of the laser field emerging from the nonlinear scattering of the valence electrons by the crystal potential contain the critical information that enables picometre-scale, real-space mapping of the valence electron structure. We used high harmonics to reconstruct images of the valence potential and electron density in crystalline magnesium fluoride and calcium fluoride with a spatial resolution of about 26 picometres. Picometre-scale imaging of valence electrons could enable direct probing of the chemical, electronic, optical and topological properties of materials.

The generation of high harmonics in solids[7–12] has led to numerous advances in strong-field condensed-matter physics. High harmonics in solids are primarily interpreted as the result of the nonlinear driving of electrons within and between bands[16–20]. High harmonics in solids are now used to probe the essential characteristics of solids, such as the band dispersion[9,21,22], the topology[13], the dynamic conductivity[23] and the arrangement of atoms in the crystal lattice[13,14]. Yet, the direct imaging of the valence electron potential and density of crystalline solids requires a description of light–matter interactions in solids within the framework of scattering[24,25], as typically used in atomic-scale diffraction microscopies[26].

It is now understood that laser fields can modify the electrostatic potential of solids and thereby be used to manipulate their electronic gaps and structure[27–30], providing ample opportunities for optical engineering of materials[31,32]. Yet, the interpretation of the interaction of a laser and crystal electrons, and the associated nonlinear emission of radiation, within the framework of scattering is more demanding. The laser fields should be sufficiently strong and fast to effectively supress the valence crystal potential so that it becomes a weak perturbation to the laser-driven motion of electrons. Ultrafast laser pulses[33], which are capable of damage-free exposure of bulk solids at fields that exceed their static dielectric strength by many orders of magnitude[7–14,23,34], could enable this possibility.

To better appreciate how a scattering regime could possibly emerge in the extreme nonlinear optics of solids, we consider the interaction of valence electrons in a crystal potential $V(r)=\sum_k V_k \mathrm{e}^{ikr}$ with a laser field $F(t)=F_0\sin(\omega_\mathrm{L}t)$. Here k and V_k denote the reciprocal lattice vectors and the Fourier components of the crystal potential, respectively, r is the spatial coordinate and i is the imaginary unit. F_0 and ω_L denote the amplitude and frequency of the laser field, respectively, and t is time. Following earlier studies of atoms and solids[27–29,35–39] and by expressing the time-dependent Schrodinger equation within the reference frame of the moving electron under the laser field, the total potential experienced by a valence electron can be formulated as[27–29]

$$U(r,t)=\sum_k V_k J_0\left(\frac{kF_0}{\omega_\mathrm{L}^2}\right)\mathrm{e}^{ikr}+\sum_{N\neq 0,\,k} V_k J_N\left(\frac{kF_0}{\omega_\mathrm{L}^2}\right)\mathrm{e}^{i(kr-N\omega_\mathrm{L}t)} \tag{1}$$

Here J_N is the Bessel function of the first kind and order N. The first sum of terms in equation (1) is time independent and describes an effective crystal potential $V_\mathrm{eff}(r,F_0)=\sum_k V_k J_0\left(\frac{kF_0}{\omega_\mathrm{L}^2}\right)\mathrm{e}^{ikr}$, the properties of which are modified by the laser field. This potential is analogous to the Krammer–Henneberger potential in atomic physics[35–39]. The rest of the terms in equation (1) describe transitions (absorption and emission of photons at harmonic energies $N\omega$ of the fundamental field) among the states of the effective crystal potential[36,37,39]. Thus, the character of the optical interaction can now be intuitively understood in the framework of the effective potential $V_\mathrm{eff}(r,F_0)$ and its quantum states under the intense laser field F_0.

The valence electrons can now be driven as quasi-free particles by an external laser[27–29] when $V_\mathrm{eff}(r,F_0)\approx 0$. For the dominant reciprocal-space vector of a crystal ($k=2\pi/d$, where d is the lattice constant), this implies that $J_0\left(\frac{kF_0}{\omega_\mathrm{L}^2}\right)\approx 0$ and suggests that for a typical

---

[1] Institut für Physik, Universität Rostock, Rostock, Germany. [2] Max-Planck-Institut für Quantenoptik, Garching, Germany. [3] Institute of Physics, Chinese Academy of Sciences, Beijing, China. [4] These authors contributed equally: H. Lakhotia, H. Y. Kim. ✉e-mail: e.goulielmakis@uni-rostock.de

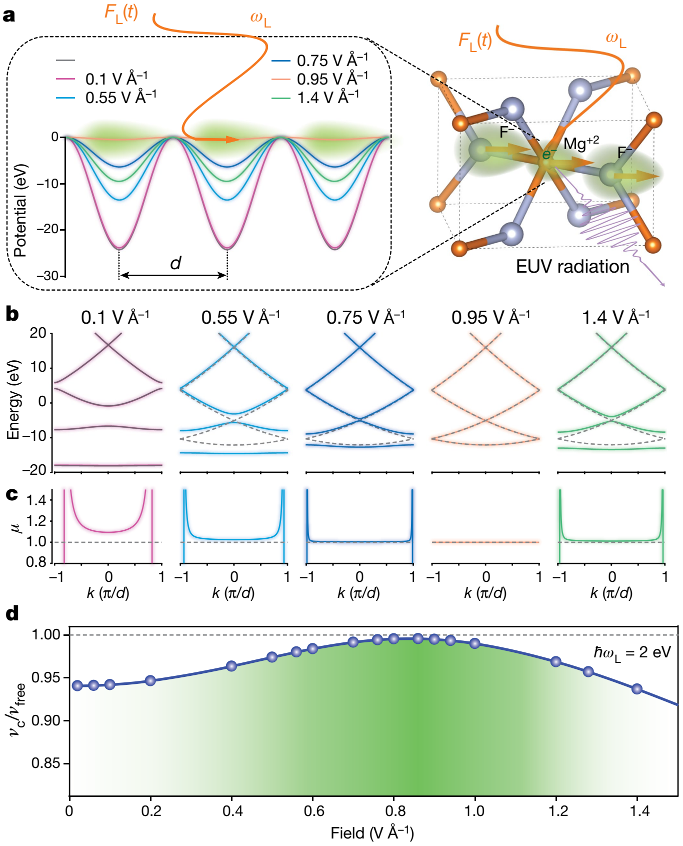


Fig. 1 | Strong-field quasi-free electron motion in a crystal. a, Effective crystal potential of MgF_2 along the [100] axis for intense optical ($\hbar\omega_L = 2$ eV) fields of increasing strength (0.1–1.4 V Å^{-1}). The field-free crystal potential is shown as a black curve. Green shaded areas indicate the valence electron cloud. EUV, extreme ultraviolet. **b**, **c**, The band structure (**b**) and reduced effective mass (**c**) as calculated for the three lowest bands and for the corresponding field strengths shown in **a**. The black curve in **b** and **c** denotes the band structure and effective mass of the undressed solid, respectively. Dashed lines in **b** and **c** represent the band dispersion and effective mass of free electrons, respectively. **d**, Ratio of the maximum of crystal ($v_c(t)$) and free ($v_{free}(t)$) electron velocities along [100] direction of MgF_2 crystal calculated by TDDFT as a function of field strength of the driving pulse with a carrier photon energy of 2 eV. The blue curve and the grey dashed line are guides for the eye.

solid ($d = 2$–7 Å) exposed to an optical field ($\hbar\omega_L \approx 2$ eV), a quasi-free electron motion will emerge for fields in the range of 0.4–1.4 V Å^{-1}, typically attainable in strong-field experiments in solids[9,11,23,34].

Figure 1a (left) shows the modification of $V_{eff}(r, F_0)$ in magnesium fluoride (MgF_2; right) along the [100] direction as calculated by an optical field ($\hbar\omega_L \approx 2$ eV) of gradually increasing F_0. At low fields, $F_0 < 0.1$ V Å^{-1}, $V_{eff}(r, F_0)$ (Fig. 1a, magenta curve) is nearly identical to that of the unperturbed crystal (Fig. 1a, black curve). The associated band structure (Fig. 1b, magenta curve) and the reduced effective mass of carriers μ in the crystal (Fig. 1c, magenta curve) hardly differ from those of the undressed solid (black curve in Fig. 1b).

At higher fields, the effective potential $V_{eff}(r, F_0)$ is notably supressed (Fig. 1a, cyan and blue curves). The band structure (Fig. 1b, cyan and blue curves) and μ (Fig. 1c, cyan and blue lines) are now gradually approaching that of the free electron (grey dashed lines in Fig. 1b, c, respectively). At the critical field strength $F_0 \approx 0.93$ V Å^{-1} for which $J_0\left(\frac{kF_0}{\omega_L^2}\right) \approx 0$, the crystal potential is totally suppressed (Fig. 1a, orange curve). Bandgaps (Fig. 1b, orange curve) are now coherently closing[27–29], while the band dispersion (Fig. 1b, orange curve) and μ (Fig. 1c, orange curve) of the carriers in the crystal virtually match those of the free electron (grey dashed curves in Fig. 1b, c, respectively). For ever higher fields, the $V_{eff}(r, F_0)$ revives (Fig. 1a, green curve), obeying the oscillatory nature of $J_0\left(\frac{kF_0}{\omega_L^2}\right)$, versus field F_0 (Fig. 1a). The reduced effective mass μ remains near that of the free electron (Fig. 1c, green line) for the best part of the Brillouin zone, but its sharp discontinuity at the edges is restored.

Figure 1a–c clearly shows that the notion of quasi-free electron motion becomes plausible over a broad range of laser fields provided that the charge carriers in the crystal do not reach the edges of the Brillouin zone to experience a Bragg reflection. This implies that for the crystal momentum $k(t) < \pi/d$ and for the optical field amplitude $F_0 < \frac{\pi\omega_L}{d}$.

First-principle, time-dependent density functional theory (TDDFT) simulations in three dimensions (Fig. 1d) on crystalline MgF_2 exposed to few-cycle pulses ($\hbar\omega_L \approx 2$ eV) with an electric field vector aligned with the [100] axis of the crystal support the above perspective (Methods). To allow an intuitive relation to Fig. 1a–c, we calculated the velocity of the carriers in the bulk crystal $v_c(t)$ (see also Extended Data Fig. 1) and compared it with that of the free electrons $v_{free}(t)$, exposed to identical waveforms and for a wide range of optical field strengths. As the reduced effective mass μ in the crystal is approximately related to the velocity ratio as $v_c(t)/v_{free}(t) \approx 1/\mu$, this calculation allows us to place the perspective of Fig. 1a–c under further scrutiny.

The ratio $v_c(t)/v_{free}(t)$, as evaluated at the maximum of $v_{free}(t)$, is shown in Fig. 1d. For weak fields, this ratio reflects the inverse of the reduced effective mass of carriers $1/\mu \approx 0.94$ around the Γ point of MgF_2. For higher field strengths, the velocity of the crystal electrons rapidly increases and reaches that of the free electron at the critical field $F_0 \approx 0.95$ V Å^{-1} (Fig. 1c, orange curve). For a further increase of the optical field $F_0 > \pi\omega_L/d$ (Fig. 1d), the simulations verify that electrons will experience Bragg reflection as manifested by the rapid drop in the carrier velocity in the crystal with respect to that of the free electrons.

Within the range of optical fields for which the crystal potential is considerably suppressed (Fig. 1a) and the band structure assumes a quasi-parabolic profile, the dynamics of the electronic wavefunction and concomitant emission of high harmonics may be treated within the framework of scattering and the motion of the electron in the potential can be described both classically[24] and quantum mechanically[25]. For MgF_2, this is further verified by comparing high-harmonic emission as described in these models with the results of the TDDFT simulations (Extended Data Fig. 2).

Within the scattering picture, the electric currents and emitted radiation are linked (Methods) to the valence periodic potential of the crystal as:

$$\frac{\partial \mathbf{J}}{\partial t} \propto N_e \sum_{\mathbf{k}} \mathbf{k} V_{\mathbf{k}} \sum_{N} J_N\left(\frac{\mathbf{k}\cdot\mathbf{F}_0}{\omega_L^2}\right) e^{iN\omega_L t} \qquad (2)$$

where N_e is the number of electrons.

Extending to three dimensions (Methods), the intensity of the harmonics that are collinearly polarized with the unit vector of the laser polarization ($\mathbf{e}_l$) is linked with the periodic potential of the crystal as:

$$I_N(F_0, \omega_L, \mathbf{e}_l) \propto \left| N_e \sum_{k_l} \tilde{V}_{k_l} k_l\, J_N\left(k_l \frac{F_0}{\omega_L^2}\right)\right|^2 \qquad (3)$$

where k_l and $\tilde{V}_{k_l}$ are the projections of the reciprocal space vectors and the Fourier components of the potential, respectively, onto the laser polarization vector. In the coordinate space (Fig. 1a), this operation represents a one-dimensional (1D) slice of the potential of the crystal $V_c(\mathbf{r})$, parallel to $\mathbf{e}_l$, passing through the expectation value of the position $\mathbf{r_0}$ of the initial electronic wavefunction within a unit cell. For crystals with a centre(s) of symmetry (C), $\mathbf{r_0}$ coincides with this centre(s). Equation (3) also implies that by measuring a set of N harmonics for various laser strengths $\mathbf{F_0}$, the Fourier components $\widetilde{V}_{k_l}$ of a 1D slice (Fig. 1a, inset) of the crystal potential can be retrieved. As the intensity of every radiated harmonic of the field I_N (equation (3)) is associated with a broad range of $\widetilde{V}_{k_l}$, the relative phase information—in contrast to linear techniques[40]—among $\widetilde{V}_{k_l}$ is not lost; it is rather embodied in the recorded intensities I_N and thus can be also retrieved.

The highest radiated photon energy E_c (cut-off) and harmonic order N_c can now be estimated using the fact that the Bessel function in equation (3) reaches a maximum when its argument equals its order, N_c:

$$E_c = N_c\omega_L \approx \frac{k_{max}F_0}{\omega_L} \qquad (4)$$

Here, k_{max} is the highest significant (cut-off) reciprocal vector of the crystal potential $V_c(\mathbf{r})$. As previously suggested[24], equation (4) represents a clear analogy between the high-harmonic emission and the Smith–Purcell effect[41], where the radiation energy is the product of velocity and spatial frequency. k_{max} is naturally associated with the valence radius r^h of the smallest atom/ion in the system, as $k_{max} \approx 2\pi/r^h$. Therefore, the cut-off law can also be also expressed as $E_c \approx \frac{2\pi F_0}{r^h\omega_L}$, suggesting that within the scattering approximation, the dimensions of the smallest atomic or ionic radii in a crystal are directly linked to the cut-off energy and thus can be probed by measuring the cut-off energy as:

$$r^h \approx \frac{2\pi F_0}{E_c\omega_L} \qquad (5)$$

Probing the ion/atomic radii in solids

In a first set of experiments, we interrogate the validity of the scattering picture by examining the possibility of probing the smallest ionic/covalent radii of atoms in solids with a sole measurement of the high-harmonic cut-off energy, as suggested by equation (5). In the experiments summarized in Fig. 2a, strong ($F_0 = 0.4$–$0.7\ \mathrm{V\,Å^{-1}}$), few-cycle pulses (duration of about 5.5 fs) carried in the visible ($\hbar\omega_L \approx 2$ eV) generated harmonics in MgF_2 and other crystalline solids (Methods). The properties of the driving pulses—including the peak electric field F_0 and centroid carrier frequency ω_L—in these experiments are accessed by attosecond streaking[42]. Representative harmonic spectra recorded in MgF_2 when the laser polarization vector is aligned with the [110] and [100] crystal axes are shown in Fig. 2b. For MgF_2, we record the cut-off energy E_c as a function of the optical field strength F_0 (Fig. 2c, black dots). We evaluate the corresponding slope of E_c versus the field strength F_0 in Fig. 2c (blue line) and derive the radius (equation (5)) as $r^h = 59 \pm 4$ pm. This result reasonably agrees with the empirical radius[43] of Mg^{+2} in MgF_2 (about 72 pm) and lies far from the corresponding radius of the much larger F^- (about 130 pm). To interrogate the character of these findings more generally, we extended the measurements to several crystalline materials, as summarized in Fig. 2d. The radii evaluated from these measurements (Fig. 2d, blue bars) once again agree well with the empirical predictions (Fig. 2d, red bars) and suggest that the scattering model is applicable in these systems.

Mapping of the crystal potential and electron density in MgF_2 and CaF_2

In a next set of experiments, we set the laser polarization parallel to a specific crystal direction by rotating the crystal (Fig. 2a) and recorded the harmonic yield I_N versus F_0 and the crystal angle (Extended Data Fig. 3). For the [110] and [100] crystal axes of MgF_2, the recorded harmonic yields are shown in Fig. 3a, b, respectively. An excellent fit (Methods) of the experimental data (red and blue curves in Fig. 3a, b, respectively) using equation (3) is obtained for all harmonics. The retrieved amplitudes and phases of $\widetilde{V}_{k_l}$ are shown in Fig. 3c, d, respectively.

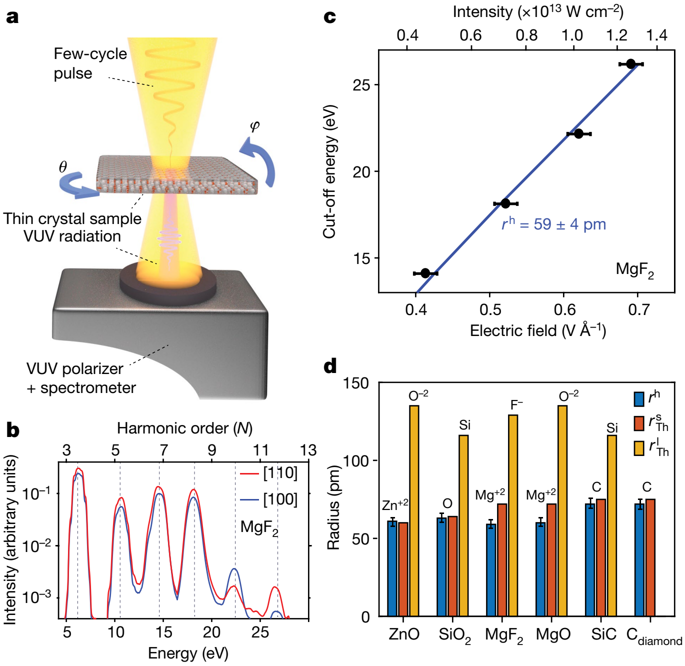


Fig. 2 | Probing of the ionic/covalent radius of atoms in solids. **a**, Experimental setup for laser picoscopy. θ and φ denote the azimuthal angle and polar angle, respectively. **b**, Harmonic spectra generated in MgF_2 when the laser polarization vector is aligned with the [110] (red curve) and [100] (blue curve) axes of the crystal. **c**, High harmonic cut-off energy E_c (black dots) versus driving electric field amplitude F_0 for the [110] axis of the crystal. The ionic radius ($r^h = 59 \pm 4$ pm) of Mg^{+2} is evaluated by the slope of the data (blue line) according to equation (5). **d**, Measured (blue bars) and empirical (red bars) ionic/covalent radii of the smallest ions/atoms (r_{Th}^{s}) in ZnO, SiO_2, MgF_2, MgO, SiC and diamond crystals. The empirical valence radii of the largest ions/atoms (r_{Th}^{l}) in each crystal (orange bars) are shown for comparison. Error bars indicate the standard deviation of the mean value from three measurements acquired under identical conditions.

The inverse Fourier transform of these data yields the reconstructed, real-space potential shown in Fig. 4a, b. A measurement (Methods) along the [110] axis provides a 1D slice of the crystal potential along a line defined by the crystal symmetry point and the laser polarization vector $\mathbf{e}_l$. In MgF_2, this implies probing of the crystal potential along F–Mg–F axis on plane (002) (Fig. 3a). Indeed, the retrieved potential slice (Fig. 4a) is composed of three consecutive valleys, which can intuitively be assigned to the F^-, Mg^{+2} and F^- ion potentials, respectively. Yet, because a second symmetry point (C_2) of MgF_2 is located on plane (001), an additional 1D potential slice (along a line defined by C_2 and the laser polarization vector $\mathbf{e}_l$) contributes to the measurement. The addition of these two 1D potential slices practically results in the duplication of the Mg^{+2} contribution on the measured potential curve of Fig. 4a. Along the [100] axis (Fig. 4b), the potential consists of a single valley that can be primarily assigned to an Mg^{+2} ion on plane (001) plus that from an Mg^{+2} ion on plane (002). A weak contribution from the spatially extended F^- ions is also anticipated (Fig. 4b, top) and results in the broadening of the Mg^{+2} peak (Fig. 4b, blue curve).

On the basis of the data in Fig. 4a, b, as well as additional data retrieved for the intermediate crystal direction [120], we reconstruct the full 'on plane' potential of MgF_2, as shown in Fig. 4c (Methods). We also evaluate the corresponding valence electron density $n(r)$ on the basis of the Thomas–Fermi approximation as $n(r) \propto V(r)^{3/2}$, shown in Fig. 4d. The potential and electron density data in Fig. 4c, d indicate that the Mg–F 'molecular pattern' exhibits a four-fold rotational symmetry, compatible with the notion of 'simultaneous' probing of crystal planes carrying symmetry points. This aspect is further supported by the

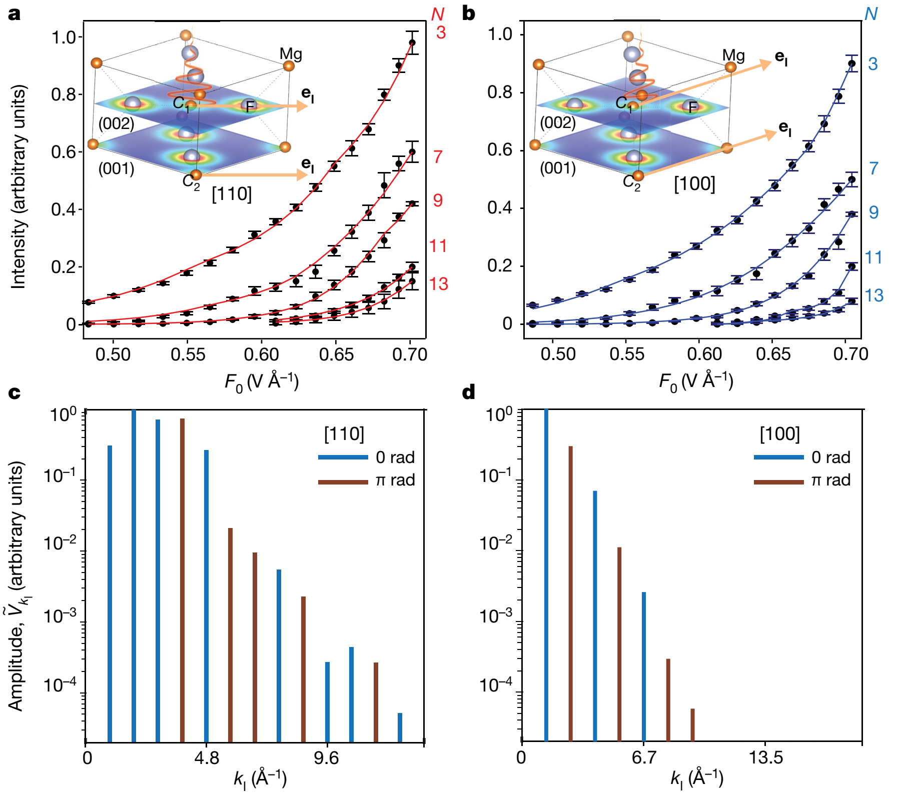


Fig. 3 | Measurement of the Fourier coefficients of the crystal potential in MgF_2. a, b, Intensity yields (black dots) of the emitted harmonics versus driving field strength F_0 as measured along the [110] (**a**) and [100] (**b**) axes of the crystal and their fittings according to equation (3), in red curves and blue curves, respectively. Error bars indicate the standard deviation of the mean value of four measurements acquired under identical conditions. Insets: MgF_2 crystal structure and DFT-calculated electron density for planes (001) and (002). The symmetry points C_1 and C_2 of the crystal are shown on the corresponding planes. The laser polarization vector (orange arrow) and the symmetry points C_1 and C_2 determine the lines along which the Fourier coefficients of potential are probed. **c, d,** Retrieved amplitudes $\widetilde{V}_{k_l}$ and relative phases (0 rad in blue and π rad in red) along the [110] (**c**) and [100] (**d**) axes of the crystal, respectively.

simulations of Fig. 4e, which show the electron density resulting from the addition of the densities of planes (001) and (002) laterally shifted so that the symmetry points (Mg^{+2}) in each plane coincide.

The reconstructed images of the valence electrons in Fig. 4c, d enable the visualization of the atomic-scale, electronic properties of MgF_2. The electron radius of Mg^{+2} can now be directly deduced from the electron density curve (Fig. 4f)—as $r^{\mathrm{h}}_{\mathrm{Mg}^{+2}} \approx 76$ pm, which matches the empirical radius of Mg^{+2} (about 72 pm) with better accuracy compared with the estimative measurements based on the cut-off method (Fig. 2c). We also evaluate the corresponding radius of F^- as $r^{\mathrm{h}}_{\mathrm{F}^-} \approx 126$ pm, which is also in agreement with the empirical data (about 130 pm).

The direct measurement of the valence electron structure in solids with picometre accuracy in these systems enables a direct comparison

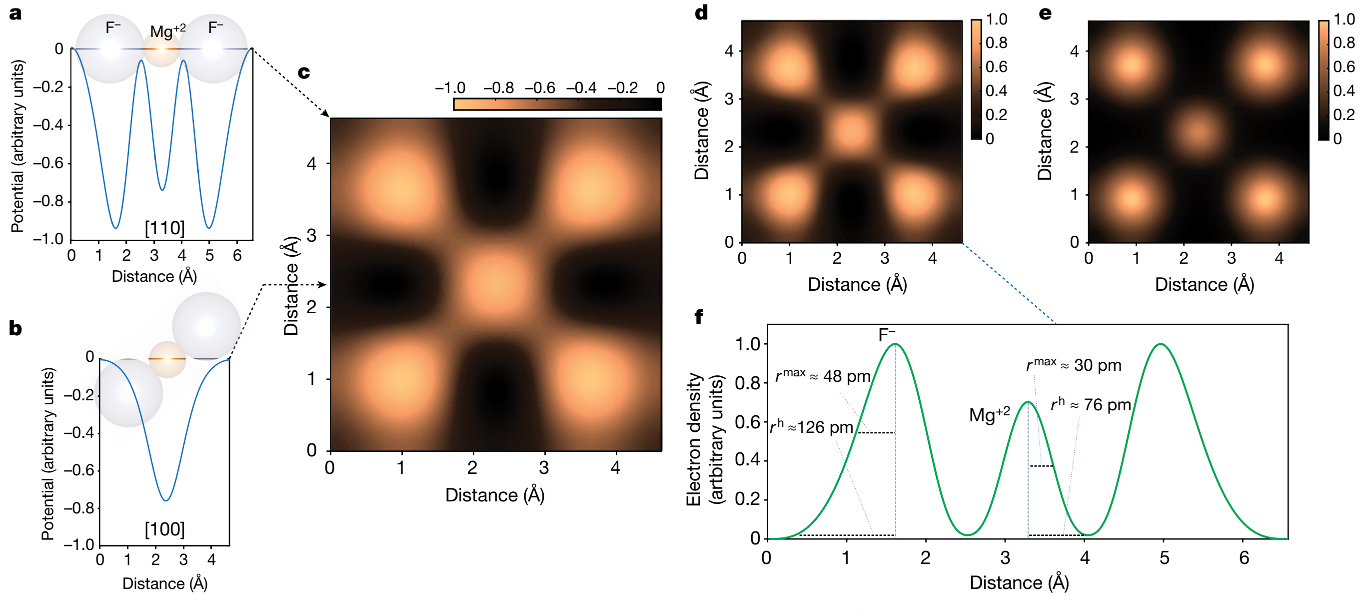


Fig. 4 | Reconstruction of the valence electron potential and density of MgF_2. a, b, Reconstructed 1D slices of the valence potential (blue curves) when the laser polarization vector is aligned with the [110] (**a**) and [100] (**b**) axes. Grey and orange spheres represent F^- and Mg^{+2} ions, respectively, as aligned along the probed line of the crystal. **c,** A reconstructed 2D slice of the valence electron potential of MgF_2. The Mg^{+2} ion is in the centre, surrounded by F^- ions. **d,** Valence electron density derived from the data in **c**. **e,** DFT-simulated electron density of MgF_2 summed over the (001) and (002) planes and shifted such that the symmetry points (C_1 and C_2 as shown in the insets of Fig. 3a, b) on both the planes coincide. **f,** Electron density (green curve) along the F–Mg–F axis of the MgF_2 crystal derived from the experimentally reconstructed valence electron potential shown in **a**. Black dashed lines indicate the evaluated ionic radii (r^{h}) and ($r^{\max}$) as defined in the text.

of experimental and rigorously calculated quantum mechanical quantities. The electron radius, which is often associated with essential properties of materials such as the polarizability and diamagnetic susceptibility, is defined as the principal maximum ($r^{\max}$) of the radial density distribution function[44]. Evaluations of $r^{\max}$ from the retrieved data for each ion in Fig. 4f yielded $r_{\mathrm{Mg}^{+2}}^{\max} \approx 30$ pm and $r_{\mathrm{F}^{-}}^{\max} \approx 48$ pm. Once again, these values closely match the theoretically calculated radii of 27 pm and 44 pm for Mg^{+2} and F^{-}, respectively[44]. We further benchmarked the capability of laser picoscopy to image the valence electronic structure by extending our experimental study to a system with a rather different crystalline structure: calcium fluoride (CaF_2, fluorite) (Methods, Extended Data Figs. 4–6).

An inspection of the reconstructed potential and/or electron density distributions in the two studied systems provides information on the nature of chemical bonding. In particular, the considerable differences in the evaluated radii of the crystal ions compared with those of neutral atoms (Mg, 167 pm; F, 41 pm; Ca, 275 pm)[45] is compatible with the electron transfer from Mg and Ca to F, (Fig. 4d, Extended Data Fig. 5e, respectively) occurring during the chemical bond formation. Moreover, in MgF_2, the weak potential (Fig. 4c) and electron density (Fig. 4d) in the interstitial space between anions and cations is compatible with the ionic character of the underlying chemical bond.

The spatial resolution attained with laser picoscopy may be directly inferred from the highest reciprocal space vectors $k_{\max}$, which are substantial in the fitting of the corresponding intensity yields in Fig. 3c, d. For example, measurements along the [100] axis of MgF_2 yielded $k_{\max} = 12.2$ Å^{-1}, suggesting a spatial resolution of about 26 pm, that is, approximately half of the Bohr radius in atomic hydrogen.

Direct imaging of valence electrons in bulk solids with picometre resolution may broaden the scope of modern, atomic-scale microscopy to include direct access to the chemical, electronic and topological properties of matter. Future experiments, including these in other spectral ranges of the driving field and detailed examination and extension of the theoretical premises of picoscopy, will be required to verify the applicability of the technique to a broader range of materials. Laser picoscopy is readily incorporable with time-resolved spectroscopies and could enable the tracking of simultaneously unfolding atomic and valence electron dynamics with picometre and attosecond resolution. It may also enable a route to the detailed understanding of the phase transition dynamics of matter.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2429-z.

1. Itatani, J. et al. Tomographic imaging of molecular orbitals. *Nature* **432**, 867–871 (2004).
2. Haessler, S. et al. Attosecond imaging of molecular electronic wavepackets. *Nat. Phys*. **6**, 200–206 (2010).
3. Villeneuve, D. M., Hockett, P., Vrakking, M. J. J. & Niikura, H. Coherent imaging of an attosecond electron wave packet. *Science* **356**, 1150–1153 (2017).
4. Smirnova, O. et al. High harmonic interferometry of multi-electron dynamics in molecules. *Nature* **460**, 972–977 (2009).
5. Baker, S. et al. Probing proton dynamics in molecules on an attosecond time scale. *Science* **312**, 424–427 (2006).
6. Kübel, M. et al. Spatiotemporal imaging of valence electron motion. *Nat. Commun*. **10**, 1042 (2019).
7. Ghimire, S. et al. Observation of high-order harmonic generation in a bulk crystal. *Nat. Phys*. **7**, 138–141 (2011).
8. Schubert, O. et al. Sub-cycle control of terahertz high-harmonic generation by dynamical Bloch oscillations. *Nat. Photon*. **8**, 119–123 (2014).
9. Luu, T. T. et al. Extreme ultraviolet high-harmonic spectroscopy of solids. *Nature* **521**, 498–502 (2015).
10. Vampa, G. et al. Linking high harmonics from gases and solids. *Nature* **522**, 462–464 (2015); corrigendum **542**, 260 (2017).
11. Ndabashimiye, G. et al. Solid-state harmonics beyond the atomic limit. *Nature* **534**, 520–523 (2016).
12. Sivis, M. et al. Tailored semiconductors for high-harmonic optoelectronics. *Science* **357**, 303–306 (2017).
13. You, Y. S., Reis, D. A. & Ghimire, S. Anisotropic high-harmonic generation in bulk crystals. *Nat. Phys*. **1**, 1–6 (2016).
14. You, Y. S., Cunningham, E., Reis, D. A. & Ghimire, S. Probing periodic potential of crystals via strong-field re-scattering. *J. Phys. At. Mol. Opt. Phys*. **51**, 114002 (2018).
15. von Hoegen, A., Mankowsky, R., Fechner, M., Först, M. & Cavalleri, A. Probing the interatomic potential of solids with strong-field nonlinear phononics. *Nature* **555**, 79–82 (2018).
16. Golde, D., Meier, T. & Koch, S. W. High harmonics generated in semiconductor nanostructures by the coupled dynamics of optical inter- and intraband excitations. *Phys. Rev. B* **77**, 075330 (2008).
17. Vampa, G., McDonald, C. R., Orlando, G., Corkum, P. B. & Brabec, T. Semiclassical analysis of high harmonic generation in bulk crystals. *Phys. Rev. B* **91**, 064302 (2015).
18. Wu, M., Ghimire, S., Reis, D. A., Schafer, K. J. & Gaarde, M. B. High-harmonic generation from Bloch electrons in solids. *Phys. Rev. A* **91**, 043839 (2015).
19. Higuchi, T., Stockman, M. I. & Hommelhoff, P. Strong-field perspective on high-harmonic radiation from bulk solids. *Phys. Rev. Lett*. **113**, 213901 (2014).
20. Kemper, A. F., Moritz, B., Freericks, J. K. & Devereaux, T. P. Theoretical description of high-order harmonic generation in solids. *New J. Phys*. **15**, 023003 (2013).
21. Vampa, G. et al. All-optical reconstruction of crystal band structure. *Phys. Rev. Lett*. **115**, 193603 (2015).
22. Lanin, A. A., Stepanov, E. A., Fedotov, A. B. & Zheltikov, A. M. Mapping the electron band structure by intraband high-harmonic generation in solids. *Optica* **4**, 516–519 (2017).
23. Garg, M. et al. Multi-petahertz electronic metrology. *Nature* **538**, 359–363 (2016).
24. Hüller, S. & Meyer-Ter-Vehn, J. High-order harmonic radiation from solid layers irradiated by subpicosecond laser pulses. *Phys. Rev. A* **48**, 3906–3909 (1993).
25. Kálmán, P. & Brabec, T. Generation of coherent hard-X-ray radiation in crystalline solids by high-intensity femtosecond laser pulses. *Phys. Rev. A* **52**, R21–R24 (1995).
26. Warren, B. E. *X-ray Diffraction* (Courier Corporation, 1990).
27. Tzoar, N. & Gersten, J. Theory of electronic band structure in intense laser fields. *Phys. Rev. B* **12**, 1132–1139 (1975).
28. Miranda, L. C. M. Energy-gap distortion in solids under intense laser fields. *Solid State Commun*. **45**, 783–785 (1983).
29. Holthaus, M. The quantum theory of an ideal superlattice responding to far-infrared laser radiation. *Z. Phys. B* **89**, 251–259 (1992).
30. Gruzdev, V. E. Photoionization rate in wide band-gap crystals. *Phys. Rev. B* **75**, 205106 (2007).
31. Wang, Y. H., Steinberg, H., Jarillo-Herrero, P. & Gedik, N. Observation of Floquet-Bloch states on the surface of a topological insulator. *Science* **342**, 453–457 (2013).
32. McIver, J. W. et al. Light-induced anomalous Hall effect in graphene. *Nat. Phys*. **16**, 38–41 (2020).
33. Brabec, T. & Krausz, F. Intense few-cycle laser fields: frontiers of nonlinear optics. *Rev. Mod. Phys*. **72**, 545–591 (2000).
34. Schultze, M. et al. Controlling dielectrics with the electric field of light. *Nature* **493**, 75–78 (2013).
35. Henneberger, W. C. Perturbation method for atoms in intense light beams. *Phys. Rev. Lett*. **21**, 838–841 (1968).
36. Gavrila, M. & Kamiński, J. Z. Free-free transitions in intense high-frequency laser fields. *Phys. Rev. Lett*. **52**, 613–616 (1984).
37. Gavrila, M. Atomic stabilization in superintense laser fields. *J. Phys. At. Mol. Opt. Phys*. **35**, R147–R193 (2002).
38. Morales, F., Richter, M., Patchkovskii, S. & Smirnova, O. Imaging the Kramers–Henneberger atom. *Proc. Natl Acad. Sci. USA* **108**, 16906–16911 (2011).
39. Medišauskas, L., Saalmann, U. & Rost, J.-M. Floquet Hamiltonian approach for dynamics in short and intense laser pulses. *J. Phys. At. Mol. Opt. Phys*. **52**, 015602 (2019).
40. Taylor, G. The phase problem. *Acta Crystallogr. D* **59**, 1881–1890 (2003).
41. Smith, S. J. & Purcell, E. M. Visible light from localized surface charges moving across a grating. *Phys. Rev*. **92**, 1069 (1953).
42. Goulielmakis, E. et al. Direct measurement of light waves. *Science* **305**, 1267–1269 (2004).
43. Shannon, R. D. Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides. *Acta Crystallogr. A* **32**, 751–767 (1976).
44. Ghosh, D. C. & Biswas, R. Theoretical calculation of absolute radii of atoms and ions. Part 2. The ionic radii. *Int. J. Mol. Sci*. **4**, 379–407 (2003).
45. Ghosh, D. C. & Biswas, R. Theoretical calculation of absolute radii of atoms and ions. Part 1. The atomic radii. *Int. J. Mol. Sci*. **3**, 87–113 (2002).

Methods

DFT and TDDFT simulations in MgF_2

The time-dependent simulations were performed using the time-dependent ab initio package (TDAP)[46,47] within the framework of DFT and TDDFT[48]. First, the laser–matter interaction was included in the Hamiltonian to simulate the effect of external laser field: $\hat{H} = \frac{1}{2m}\left(\hbar, \hat{k}, -, \frac{e}{c}, \hat{A}, (t)\right)^2 + \hat{V}_c$, where m, e and k are the electron mass, charge and momentum, respectively, V_c is the periodic potential of the crystal and $\hat{A}(t) = -c\int \hat{E}(t)\mathrm{d}t$ (ref. [49]) is the vector potential of the driving pulse. c is the speed of light. The time-dependent Kohn–Sham equation (TDKS) was propagated in real time, where the propagator operator is expressed within the Crank–Nicolson scheme[46,50]. Then, the TDKS orbitals $\psi_{G,k,\sigma}(t)$ as well as the time-dependent charge density $\rho(r,t)$ were obtained at each timestep. The time-dependent electron velocity $v_{G,k,\sigma}(t)$ was calculated from the TDKS orbitals:

$$v_{G,k,\sigma}(t) = -i\sum_i \left[\psi_{i,G,k,\sigma}(t)|\nabla|\psi_{i,G,k,\sigma}(t) - \psi_{i,G,k,\sigma}(t)|\nabla|\psi_{i,G,k,\sigma}(t)^*\right] \quad (6)$$

where i labels the state index and G, k and σ are the planewave-basis, k point and spin index, respectively. We derive the total electron velocity as:

$$v(t) = \frac{1}{V_{\mathrm{cell}}}\sum_{G,k,\sigma} v_{G,k,\sigma}(t) \quad (7)$$

where V_{cell} is the volume of the unit cell. More details about the algorithm can be find in refs. [46,47]. The velocity of the free electron was trivially obtained from the vector potential $A(t)$ as: $v(t)_{\mathrm{free}} = \hbar A(t)/m$, where m is the free electron mass. In the calculations, we considered laser pulses with the characteristics of those used in the experiments. We used norm-conserving pseudopotentials with the Perdew–Burke–Ernzerhof functional[51]. To reduce the computation time, we opted for the numerical atomic orbitals as well as an auxiliary real-space grid equivalent to a planewave cut-off of 150 Rydberg. The k-points sampling was $6 \times 6 \times 9$. The evolution of the system was calculated by self-consistently propagating the electron density and the results are convergent with timesteps from 2 to 20 as.

To obtain the valence electron density and electrical potential, the SIESTA DFT package[52] was used with the PBE[51] functional and numerical atomic orbitals basis set. For sampling, $12 \times 12 \times 18$ points in 3D k-space were used, and a density matrix was calculated to obtain the electron density. The reduced effective mass in MgF_2 was estimated on the basis of two valence and two conduction bands within the same DFT code.

High-harmonic generation in the semi-classical limit

The emission of harmonics from a solid driven by an intense laser field $\mathbf{F}_\mathrm{L}(t) = F_0\sin(\omega_\mathrm{L}t)\mathbf{e}_\mathbf{l}$ is associated with rate of change of the induced current in its bulk. Under conditions for which the crystal potential is softened by the intense field and the corresponding band structure, in turn, becomes a quasi-parabolic profile as presented in the main text, the kinematics of the electrons can be treated semi-classically by introducing the limit[53] $\nabla V_\mathrm{c}(\mathbf{r}(t)) \approx \nabla V_\mathrm{c}(\langle\mathbf{r}(t)\rangle)$. In this limit, the current variation in time is governed by the classical equation of motion according to the Ehrenfest theorem.

$$\frac{\partial}{\partial t}\mathbf{J}(t) \propto -N_\mathrm{e}\nabla V_\mathrm{c}(\mathbf{r}(t)) \approx -N_\mathrm{e}\nabla V_\mathrm{c}(\langle\mathbf{r}(t)\rangle) \quad (8)$$

$V_\mathrm{c}(\langle\mathbf{r}(t)\rangle)$ in this case represents the potential of N_e valence electrons in the crystal, and $\langle\mathbf{r}(t)\rangle$ stands for the classical expectation value of the position of the wavepacket within a unit cell of a crystal following the electric field of the laser as $\langle\mathbf{r}(t)\rangle = \left(\mathbf{r}_0 + \frac{F_0}{\omega_\mathrm{L}^2}\sin(\omega_\mathrm{L}t)\right)$. $\mathbf{r}_0$ is the expectation value of the initial position of the electron wavefunction within a unit cell. The periodic potential can now be expressed in terms of its Fourier coefficients $V_\mathrm{c}(\mathbf{r}) = \sum_\mathbf{k} V_\mathbf{k}\mathrm{e}^{-i(\mathbf{k}\cdot\mathbf{r})}$

$$\frac{\partial}{\partial t}\mathbf{J}(t) \propto -N_\mathrm{e}\nabla\sum_\mathbf{k} V_\mathbf{k}\mathrm{e}^{(-i\mathbf{k}\cdot\langle\mathbf{r}(t)\rangle)} = iN_\mathrm{e}\sum_{\boldsymbol{k}} \mathbf{k}V_\mathbf{k}\mathrm{e}^{\left(-i\mathbf{k}\cdot\left(\mathbf{r}_0 + \frac{\mathbf{F}_0}{\omega_\mathrm{L}^2}\sin(\omega_\mathrm{L}t)\right)\right)} \quad (9)$$

The exponential term in the above equation can be further expanded using the Jacobi–Anger expansion. The real part of $\frac{\partial}{\partial t}\mathbf{J}(t)$ can be decomposed into odd and even terms such as:

$$\frac{\partial}{\partial t}\mathbf{J}(t) = \frac{\partial}{\partial t}\mathbf{J}^{\mathrm{odd}}(t) + \frac{\partial}{\partial t}\mathbf{J}^{\mathrm{even}}(t)$$

where

$$\frac{\partial}{\partial t}\mathbf{J}^{\mathrm{odd}}(t) \propto N_\mathrm{e}\sum_\mathbf{k}\mathbf{k}\mathrm{Re}(V_\mathbf{k})\sum_{n=1}^{\infty} J_{2n-1}\left(\frac{\mathbf{k}\cdot\mathbf{F}_0}{\omega_\mathrm{L}^2}\right)\sin[(2n-1)\omega_\mathrm{L}t] \quad (10)$$

$$\frac{\partial}{\partial t}\mathbf{J}^{\mathrm{even}}(t) \propto N_\mathrm{e}\sum_k\mathbf{k}\mathrm{Im}(V_\mathbf{k})\sum_{n=1}^{\infty} J_{2n}\left(\frac{\mathbf{k}\cdot\mathbf{F}_0}{\omega_\mathrm{L}^2}\right)\cos[(2n)\omega_\mathrm{L}t] \quad (11)$$

where J_N is the Bessel function of the first kind and order N, and $\mathrm{Re}(V_\mathbf{k})$ and $\mathrm{Im}(V_\mathbf{k})$ are the real and the imaginary parts of $V_\mathbf{k}$. The intensity I_N of a radiated harmonic of order N as a function of the field strength F_0 and driving frequency ω_L is given by a square modulus of the Fourier transform of the rate of change of the total current. In the experiments presented here, only centrosymmetric crystals were used ($\mathrm{Im}(V_\mathbf{k}) = 0$), hence only odd harmonics are relevant. The intensity I_N of the odd harmonics can be further decomposed into components parallel ($\mathbf{e}_\mathbf{l}$) and perpendicular ($\mathbf{e}_\perp$) to the laser polarization vector:

$$I_N(F_0, \omega_\mathrm{L}, \mathbf{e}_\mathbf{l}) \propto \left|N_\mathrm{e}\sum_\mathbf{k}\mathbf{k}\cdot\mathbf{e}_\mathbf{l}\ V_\mathbf{k}\ J_N\left(\mathbf{k}\cdot\mathbf{e}_\mathbf{l}\frac{F_0}{\omega_\mathrm{L}^2}\right)\right|^2 \quad (12)$$

$$I_N(F_0, \omega_\mathrm{L}, \mathbf{e}_\perp) \propto \left|N_\mathrm{e}\sum_\mathbf{k}\mathbf{k}\cdot\mathbf{e}_\perp V_\mathbf{k}\,J_N\left(\mathbf{k}\cdot\mathbf{e}_\mathbf{l}\frac{F_0}{\omega_\mathrm{L}^2}\right)\right|^2 \quad (13)$$

Notably, the results of equations (12) and (13) closely match those of previous studies that treated the problem classically[24] or quantum mechanically[25].

All the factors except $V_\mathbf{k}$ inside the summation in equation (12) depend on $\mathbf{k}$ through $\mathbf{k}\cdot\mathbf{e}_\mathbf{l}$. This fact is used to split the summation into two parts, one in the direction of laser polarization ($\mathbf{e}_\mathbf{l}$) and other perpendicular to it ($\mathbf{e}_\perp$). The summation of the Fourier coefficients of the potential $V_\mathbf{k}$ in the perpendicular direction ($\mathbf{e}_\perp$) is given as:

$$\widetilde{V}_{k_\mathrm{l}} = \sum_{\mathbf{k}_\perp} V_{\mathbf{k}_\mathbf{l}+\mathbf{k}_\perp} \quad (14)$$

where $\mathbf{k} = \mathbf{k}_\mathbf{l} + \mathbf{k}_\perp$ such that $k_\mathrm{l} = \mathbf{k}\cdot\mathbf{e}_\mathbf{l}$ and $k_\perp = \boldsymbol{k}\cdot\boldsymbol{e}_\perp$ are the projection of reciprocal space vector in the direction parallel and perpendicular to laser polarization, respectively. According to the Fourier slice theorem[54], the Fourier transform of $\widetilde{V}_{k_\mathrm{l}}$ represents a 1D slice of the potential in the $\mathbf{e}_\mathbf{l}$ direction, passing through the origin $\mathbf{r}_0$. Hence, the emitted radiation $I_N(F_0, \omega_\mathrm{L}, \mathbf{e}_\mathbf{l})$ is associated with the motion of the electron along this slice. In this case, equation (12) is reduced to a scalar form:

$$I_N(F_0, \omega_\mathrm{L}, \mathbf{e}_\mathbf{l}) \propto \left|N_\mathrm{e}\sum_{k_\mathrm{l}}\widetilde{V}_{k_\mathrm{l}}k_\mathrm{l}J_N\left(k_\mathrm{l}\frac{F_0}{\omega_\mathrm{L}^2}\right)\right|^2 \quad (15)$$

An important implication of the scattering approximation as suggested by the derivations summarized in equation (15) or equation (3) is that, as the valence electron cloud, in the entire volume of the unit cell, can be considered free to move under the driving field, its dynamics can be described by the temporal evolution of the expectation value of the wavefunction $\mathbf{r}(t)$ (single-point dynamics). The weak perturbation from the crystal potential to the electron-cloud motion induces coherent currents and consequentially harmonic radiation. Harmonic yields calculated by TDDFT and the scattering model are in a good agreement (Extended Data Fig. 2).

Experimental

MgF_2, MgO, CaF_2, SiO_2, ZnO and diamond crystals with thickness of about 2 μm or less were placed in a vacuum chamber and were exposed to few-cycle (about 5.5 fs) pulses carried at about 2 eV. The pulses were produced by a second-generation light-field synthesizer[55]. The field strength F_0 of the laser pulses on the sample was varied by a precisely adjustable aperture. The detailed field waveform (including the amplitude F_0 and the carrier frequency ω_L) of the pulses was measured using attosecond streaking[42]. The vacuum ultraviolet (VUV) radiation emerging from the sample was polarization-filtered by reflection off two rhodium-coated concave mirrors placed at a quasi-grazing incidence (about 78°) as well as a flat-field VUV grating (about 75°). The intensity of the perpendicular polarized ($\mathbf{e}_\perp$) component of the emitted high harmonics was suppressed by a factor of about 20 compared with the parallel polarized component ($\mathbf{e}_\parallel$). The grating was also used to disperse the harmonic spectrum on a microchannel plate-phosphor screen detector. A high-dynamic-range charge-coupled device camera recorded images of high-harmonic spectra versus the driving field strength and the crystal angle. Cut-offs in all measurements were defined as the last harmonic energy detectable by our detection system. Owing to a strong contamination of the fifth harmonic signal by second-order diffractions of the grating, its intensity dependence has been omitted in the potential reconstruction of MgF_2.

Reconstructing 1D slices of the crystal potential

The intensity yields I_N of the recorded harmonics versus driving field amplitude F_0 were used to retrieve the amplitudes and phases of the Fourier coefficients of the 1D slice of the crystal potential as described by equation (15) or equation (3). As in this first study we are interested in the relative amplitudes and phases among these coefficients, the intensity yields of all harmonics were normalized to unity. This step is essential in the reconstruction process as it prevents artefacts that relate to the accurate knowledge of the relative yields of harmonics. The latter can be affected by the transmission of the specimen and the intensity calibration of the detection system—especially over the extended spectral ranges of study.

A least-squares fitting algorithm (Levenberg–Marquardt) was used to fit the experimental data within MATLAB R2016b. The fitting converged rapidly and yielded a regression better than 3%. The linear slices of the potential resulting from the inverse Fourier transform of the projected coefficients are plotted in arbitrary amplitude units.

Strictly, equation (15) or equation (3) are accurate for a monochromatic light field. In the case of a pulsed driving field, the equation of motion for the total current can expressed as

$$\frac{\partial}{\partial t}\mathbf{J}(t) \propto iN_e \sum_{\mathbf{k}} \mathbf{k} V_{\mathbf{k}} e^{\left(-i\mathbf{k}\cdot\left(\frac{\mathbf{F_L}(t)}{\omega_L^2}\right)\right)} \tag{16}$$

where $\mathbf{F_L}(t)$ denotes the field waveform of the driving pulse. Harmonic spectra calculated directly by equation (16) are shown in Extended Data Fig. 2b. A complete reconstruction of the Fourier coefficients can also be achieved using the above formula with $\mathbf{F_L}(t)$ being the experimentally measured electric field waveform. Yet for electric field strengths up to about 1.2 V Å^{-1}, the analytical form remains accurate under an adjustment of the field amplitude F_{0N} used in the reconstruction of the intensities for each harmonic N as

$$F_{0N} = F_0\frac{\sqrt{N\pi}}{2}\mathrm{erf}\left(\frac{1}{\sqrt{N}}\right) \tag{17}$$

where F_0 is the peak field strength of the driving pulse. In this case, the use of the quasi-analytic formulas in the main text allows a broader applicability of the technique proposed and implemented here and can also be applied in experimental facilities where a field-resolved characterization of light waveforms is not currently available.

Scattering picture and crystallography

For crystals with the rutile-structure MgF_2 in our study, the expectation value of the initial wavefunction coincides with the symmetry points of the crystal, which are the Mg^{+2} ions (marked as C_1 and C_2) on planes (001) and (002) (Fig. 3a, b, insets). As a result, when the polarization vector of the laser is, for example aligned with the [100] axis of MgF_2, the 1D slice of the potential probed will be on a line defined by the symmetry point (C_1) and laser polarization vector plus a 1D slice of the potential on a line defined by the symmetry point (C_2) and the laser polarization vector. Correspondingly, when the 2D potential on a plane is reconstructed, we anticipate that it will represent the addition of the two parallel planes on which the symmetry points lie, that is, (001) and (002), shifted so that the symmetry centres coincide. This perspective is supported by the experimental results for the potential in Fig. 4c, which indeed represent the addition of the potentials on the two planes. For the CaF_2 experiments, the symmetry point of the crystal (marked as C) lies exclusively on the (002) plane as shown in Extended Data Fig. 5a. As a result, laser picoscopy probes only a single plane (Extended Data Fig. 5e).

Reconstructing the potential on a plane

The reconstruction of the potential on a plane requires information on the properties of the lattice. The MgF_2 crystal has a square lattice when seen from the c axis, the axis through which the laser impinges on the crystal in our experiments (Fig. 3a, b, insets). Although this information can be acquired from X-ray crystallography, laser picoscopy is independent from this information. Indeed, the lattice symmetries can be directly inferred by the angular dependence of the high-harmonic intensity yield as a function of the rotation of the crystal with respect to the c axis (Extended Data Fig. 4). This feature of high harmonics has been repeatedly demonstrated in previous studies[7,8,13]. The 90° symmetry of these data directly suggests a square lattice.

A 2D potential slice $U(x, y)$ of a unit cell with a lattice constant d can be expanded in Fourier series as: $U(x,y) = \sum_{lm\in\mathbb{Z}} u_{lm} e^{i\frac{2\pi}{d}lx} e^{i\frac{2\pi}{d}my}$, where u_{lm} is the 2D Fourier coefficient of index l, and m and $\mathbb{Z}$ denote integer numbers. To reconstruct a 2D picture of the potential, we need to identify u_{lm}. In our experiments, the Fourier coefficients of the 1D slice of the potential ($\widetilde{V}_k$) along the characteristic crystal directions [100], [110] and [120] were first reconstructed. As described earlier, these Fourier coefficients are the projections of u_{lm} along the respective crystal directions. This fact is used to create a system of linear equations, which are in turn solved to obtain the amplitudes and phases of u_{lm}. Importantly, the symmetries of crystals give rise to numerous linear constraints, which reduce the number of unique unknowns of u_{lm} and simplify the problem dramatically. In this study, the system of equations was solved by a standard linear least square fitting method. The accuracy of the reconstruction depends on the number of unique 1D slices used to derive the 2D potential. To plot the 2D slice of the potential (Fig. 4c, Extended Data Fig. 5d), we kept all Fourier coefficients up to the seventh order implying a resolution of about 90 pm. Yet a more advanced implementation of this approach using several 1D potential slices acquired at different angles can restore the full resolution, which is currently available in the 1D measurements of the potentials as shown in the main text (about 26 pm).

Although in the current implementation of the technique the potentials in each slice are retrieved independently, two essential features of our approach allow their accurate combination to create 2D slices—and eventually 3D images—of the structure. The first is related to the fact that all reconstructed linear slices are centred around the same symmetry point of the crystal, or alternatively, the symmetry centre (expectation value of the position) of the crystal wavefunction within a unit cell. This implies that arbitrary phase shifts of the 1D potentials, before combining the data in a 2D picture, are not required. Second, although the 1D potential slices are reconstructed separately by the fitting of the normalized intensity harmonic yields, their relative amplitude can still be calibrated based on the variation of the harmonic yield in each direction. In the current implementation, the intensity yield of the lowest harmonic (third) was used to this end.

Reconstruction of valence potential and electron density of CaF_2

CaF_2 is best represented[56] by an expanded fcc lattice, composed of alternating layers of Ca and F atoms, as seen from the *c* axis (Extended Data Fig. 5a). A symmetry point (*C* in Extended Data Fig. 5a) of the crystal lies at the centre of the (002) (Ca plane). Therefore, we anticipate that in this crystal, laser picoscopy will be probing a single crystal plane—the Ca plane.

In the experiments, the pulses impinge on crystalline CaF_2 along the *c* axis (Extended Data Fig. 5a, orange curve). One-dimensional potential slices reconstructed by recording the intensity yield of harmonics versus the field strength (Extended Data Fig. 4) along the [110] and [100] axes are shown in Extended Data Fig. 5b, c, respectively. The derived 2D potential and electron density slices (for the plane (002)) are shown in Extended Data Fig. 5d, e, respectively as well as in Extended Data Fig. 6.

Beyond the anticipated dominance of Ca^{+2} on plane (002), a notable contribution from F^- ions (Extended Data Fig. 5b, d, e) centred on the (004) plane is also observed. This can be attributed to the extended size of the ionic radius of fluorine compared with the distance between the (004) and (002) planes in CaF_2, as verified by DFT simulations (Extended Data Fig. 5f), and is suggestive of the high dynamic range provided by laser picoscopy. The radius, $r^{\max}$ for Ca^{2+} as evaluated by the experimentally derived electron density is $r^{\max}_{Ca^{+2}} \approx 50$ pm, in agreement with the theoretical prediction[44] of about 54 pm.

Data availability

The datasets generated and/or analysed during this study are available from the corresponding authors on reasonable request.

Code availability

The analysis codes that support the findings of the study are available from the corresponding authors on reasonable request.

46. Meng, S. & Kaxiras, E. Real-time, local basis-set implementation of time-dependent density functional theory for excited state dynamics simulations. *J. Chem. Phys*. **129**, 054110 (2008).
47. Lian, C., Hu, S.-Q., Guan, M.-X. & Meng, S. Momentum-resolved TDDFT algorithm in atomic basis for real time tracking of electronic excitation. *J. Chem. Phys*. **149**, 154104 (2018).
48. Runge, E. & Gross, E. K. U. Density-functional theory for time-dependent systems. *Phys. Rev. Lett*. **52**, 997–1000 (1984).
49. Bertsch, G. F., Iwata, J.-I., Rubio, A. & Yabana, K. Real-space, real-time method for the dielectric function. *Phys. Rev. B* **62**, 7998–8002 (2000).
50. Castro, A., Marques, M. A. L. & Rubio, A. Propagators for the time-dependent Kohn–Sham equations. *J. Chem. Phys*. **121**, 3425–3433 (2004).
51. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett*. **77**, 3865–3868 (1996).
52. Soler, J. M. et al. The SIESTA method for ab initio order-*N* materials simulation. *J. Phys. Condens. Matter* **14**, 2745–2779 (2002).
53. Longhi, S., Horsley, S. A. R. & Della Valle, G. Scattering of accelerated wave packets. *Phys. Rev. A* **97**, 032122 (2018).
54. Kak, A. C. & Slaney, M. *Principles of Computerized Tomographic Imaging* (Society for Industrial and Applied Mathematics, 2011).
55. Hassan, M. T. et al. Optical attosecond pulses and tracking the nonlinear response of bound electrons. *Nature* **530**, 66–70 (2016).
56. CaF_2 crystal structure: datasheet from PAULING FILE Multinaries Edition – 2012 *SpringerMaterials* https://materials.springer.com/isp/crystallographic/docs/sd_0378096 (2016).

Acknowledgements This work was supported by a European Research Council grant (Attoelectronics-258501), the Deutsche Forschungsgemeinschaft Cluster of Excellence, the Munich Centre for Advanced Photonics and the Max Planck Society.

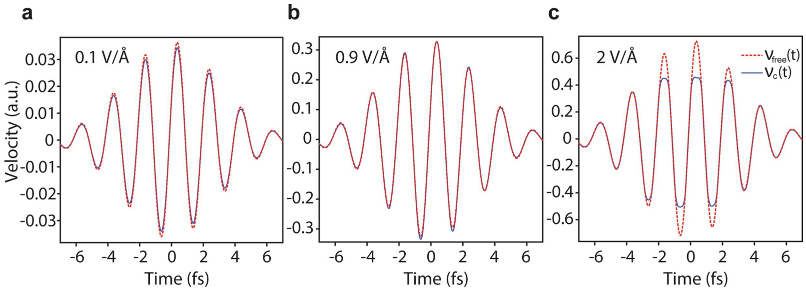
Author contributions E.G. conceived and supervised the project. H.L., H.Y.K. and M.Z. performed the experiments and analysed the experimental data. H.Y.K. and H.L. performed the theoretical modelling and calculations. S.H and S.M. conducted the DFT and TDDFT modelling. E.G., H.L. and H.Y.K. interpreted the experimental data and contributed to the preparation of the manuscript. These authors contributed equally: H. Lakhotia, H. Y. Kim.

Competing interests The authors declare no competing interests.

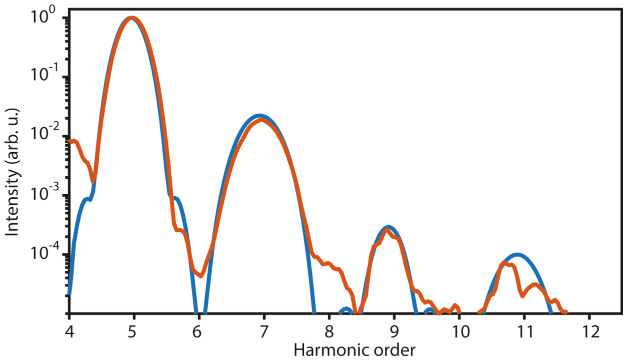
Additional information

Correspondence and requests for materials should be addressed to E.G.

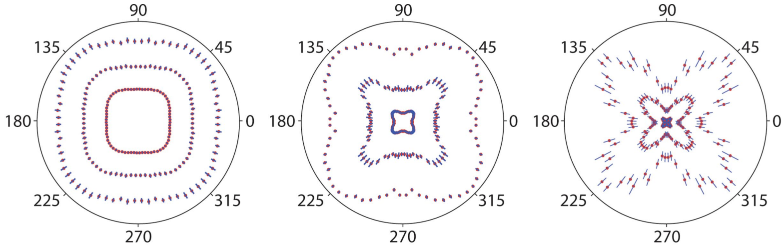
Peer review information *Nature* thanks Thomas Brabec, Michael Sentef, Andre Staudte and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



Extended Data Fig. 1 | Strong field-driven electron dynamics in MgF_2 ($\hbar\omega_L = 2eV$). a–c, Comparison of crystal (v_c; blue curves) and free (v_{free}; red dashed curves) electron velocities along the [100] direction of an MgF_2 crystal as calculated by TDDFT for laser field strengths F_0 of 0.1 V Å^{-1} (**a**), 0.9 V Å^{-1} (**b**) and 2.0 V Å^{-1}, and carrier at an energy of $\hbar\omega_L = 2eV$.



Extended Data Fig. 2 | High-harmonic generation in MgF$_2$ (theory). High-harmonic spectra calculated by TDDFT simulations (red curve) and by use of the scattering model (blue curve) for laser parameters ($\hbar\omega_L = 2$eV and $F_0 = 0.9$ V Å^{-1}) and crystal orientation settings as quoted in Fig. 1d.

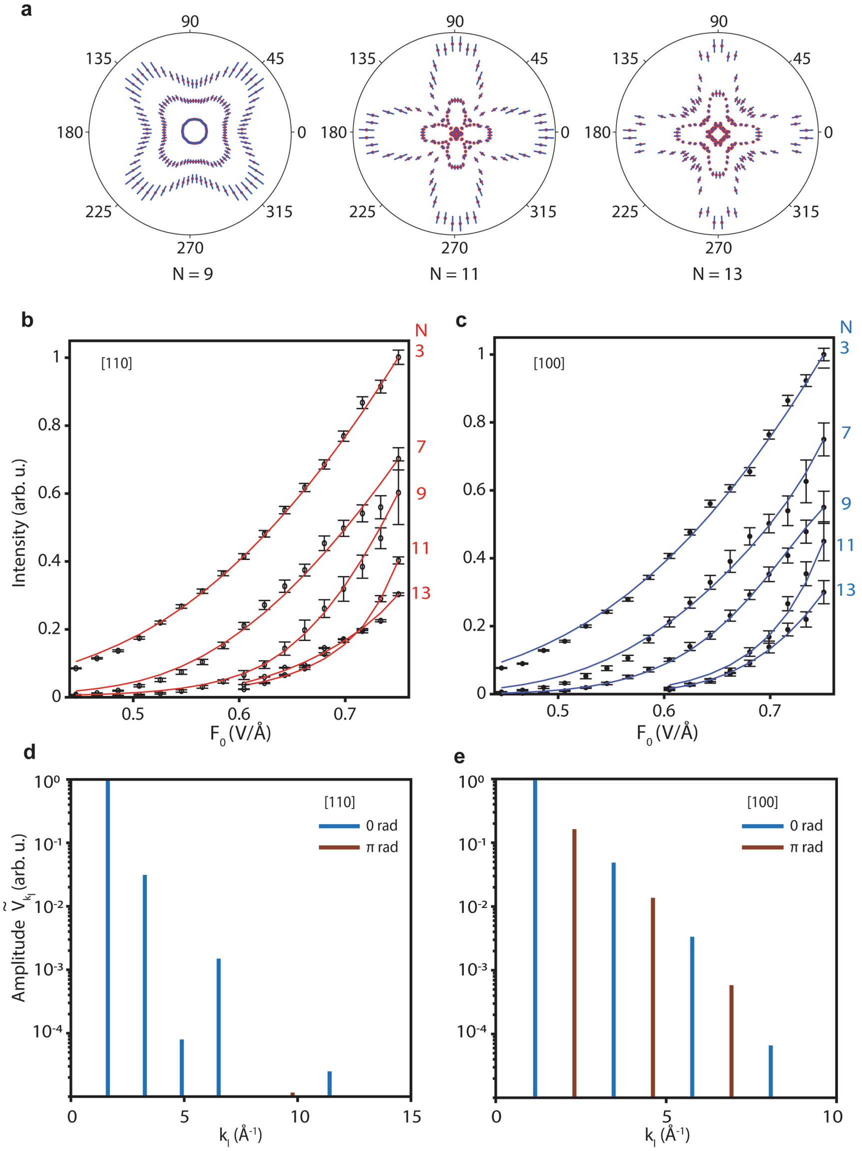


N = 3

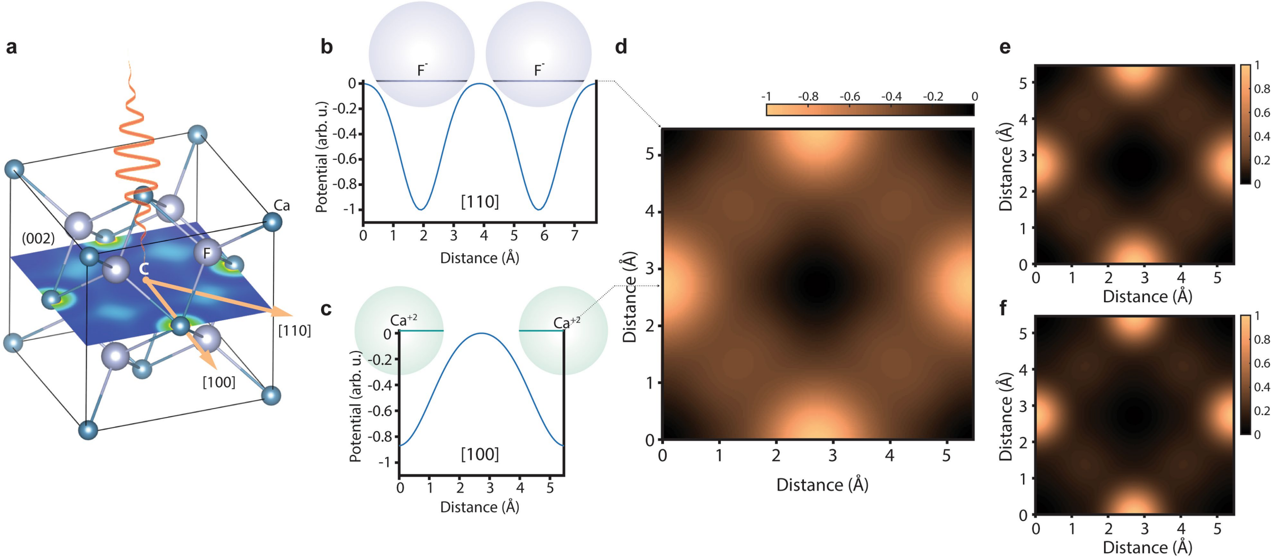
N = 9

N = 13

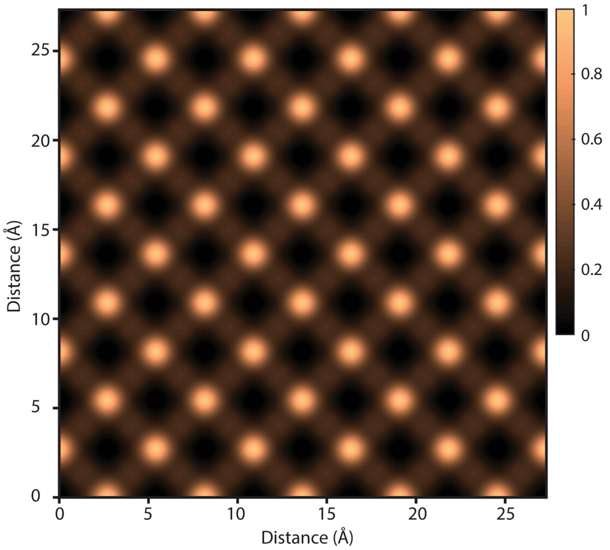
Extended Data Fig. 3 | Crystal orientation dependence of high-harmonic generation in MgF$_2$. The intensity of the third, ninth and thirteenth harmonics measured as a function of the crystal angle at field strengths (F_0 = 0.58, 0.65 and 0.7 V Å^{-1}) of the driving pulse. The rotation of the crystal is performed with respect to the *c* axis. The azimuthal angle represents the orientation of the crystal with respect to the laser polarization and the radius represents the harmonic yield. The four-fold symmetry of the crystal suggests a square lattice. Error bars in the measured data indicate the standard deviation of the mean value from four measurements acquired under identical conditions.



Extended Data Fig. 4 | Laser picoscopy in CaF_2. **a**, Intensity yields of representative harmonics (N = 9, 11 and 13) in CaF_2 measured as a function of the crystal rotation angle with respect to the *c* axis and for three representative driving field strengths (F_0 = 0.58, 0.65 and 0.7 V Å^{-1}). **b**, **c**, Intensity yields (black dots) of harmonics versus field strengths measured along the [110] (**b**) and [100] (**c**) axes of the crystal. The red and blue curves are the fitting of the intensity yields according to equation (18) or equation (3). Error bars in **a**–**c** indicate the standard deviation of the mean value from three measurements acquired under identical conditions. **d**, **e**, Retrieved amplitudes $\widetilde{V}_{k_l}$ and their relative phases (0 rad in blue and π rad in red) along the [110] (**d**) and [100] (**e**) axes of the crystal.



Extended Data Fig. 5 | Reconstruction of the valence electron potential and density of CaF_2. a, Crystal structure of CaF_2. The laser pulse (orange curve) impinges on the crystal along the *c* axis. The potential is probed along lines determined by laser polarization vectors (orange arrows) and the symmetry point *C*. **b**, **c**, Reconstructed 1D slices of the valence potential (blue curves) when the laser polarization vector is aligned with the [110] (**b**) and [100] (**c**) axes. Grey and cyan spheres represent F^- and Ca^{2+}, respectively, as aligned along the measurement line. **d**, Reconstructed 2D slice of the valence electron potential of CaF_2 on the (002) plane. Bright spots represent Ca^{+2} ions and the light broad spots represent F^- ions. **e**, Valence electron density evaluated from the data in **d**. **f**, DFT-calculated valence electron density of CaF_2 on the (002) plane.



Extended Data Fig. 6 | Electron density of CaF_2 extended over multiple unit cells. Bright dots correspond to Ca^{+2} ions centred on (002) plane while the light dots correspond to F^- ions centred on (004) plane but penetrating into the (002) plane.

Observation of branched flow of light

https://doi.org/10.1038/s41586-020-2376-8

Received: 26 September 2019

Accepted: 16 March 2020

Anatoly Patsyk[1,4], Uri Sivan[1,2], Mordechai Segev[1,2]✉ & Miguel A. Bandres[3,4]✉

When waves propagate through a weak disordered potential with correlation length larger than the wavelength, they form channels (branches) of enhanced intensity that keep dividing as the waves propagate[1]. This fundamental wave phenomenon is known as branched flow. It was first observed for electrons[1–6] and for microwave cavities[7,8], and it is generally expected for waves with vastly different wavelengths, for example, branched flow has been suggested as a focusing mechanism for ocean waves[9–11], and was suggested to occur also in sound waves[12] and ultrarelativistic electrons in graphene[13]. Branched flow may act as a trigger for the formation of extreme nonlinear events[14–17] and as a channel through which energy is transmitted in a scattering medium[18]. Here we present the experimental observation of the branched flow of light. We show that, as light propagates inside a thin soap membrane, smooth thickness variations in the film act as a correlated disordered potential, focusing the light into filaments that display the features of branched flow: scaling of the distance to the first branching point and the probability distribution of the intensity. We find that, counterintuitively, despite the random variations in the medium and the linear nature of the effect, the filaments remain collimated throughout their paths. Bringing branched flow to the field of optics, with its full arsenal of tools, opens the door to the investigation of a plethora of new ideas such as branched flow in nonlinear media, in curved space or in active systems with gain. Furthermore, the labile nature of soap films leads to a regime in which the branched flow of light interacts and affects the underlying disorder through radiation pressure and gradient force.

Waves propagating through a weak disordered potential with correlation length larger than the wavelength produce surprisingly long narrow filaments (branches)[1]. Instead of producing completely random speckle patterns, the slowly varying disordered potential gives rise to focused filaments that divide to form a pattern resembling the branches of a tree. This phenomenon is called branched flow. The underlying mechanism has been traced to deflection of rays by weak correlated variations in the potential, leading to caustics[19,20]. Formally, these caustics reflect foldings of the Lagrangian manifold in phase space[21], corresponding to the concentration of rays and high field intensity along specific lines in two dimensions or over surfaces in three dimensions. Although the nature of branched flow is linear, the high field intensity may trigger additional phenomena such as nonlinear waves (such as breather and nonlinear rogue waves)[15–17]. Branched flow is now understood to be a ubiquitous wave phenomenon, but has never been observed in optics.

Here we present the experimental observation of optical branched flow. Our experiments are carried out in thin liquid soap films (Fig. 1a, b), where the weak random correlated potential arises from naturally occurring variations in film thickness[22]. We show, in experiments, that the statistical distribution of branch intensities has a heavy tail, and that the distance from the launch point to the first branching point satisfies a scaling law that depends solely on the optical potential strength and its correlation length[23].

The experimental setting for observing branched flow in liquid soap films is shown in Fig. 1a. A soap membrane consists of a thin layer of liquid stabilized by two layers of surfactant molecules (Fig. 1b and Supplementary Fig. 1). The total thickness may vary between around 5 nm ('black film') and several micrometres, with large, naturally occurring, intra-membrane thickness variations caused by the non-uniform density of surfactant molecules. These smooth thickness variations lead to variations in the effective index of refraction for light propagating within the membrane (Fig. 1c and Supplementary Fig. 3). For thick membranes, these variations in refraction index are small but when the thickness approaches one to two wavelengths, the variations become substantial and deflect light effectively.

To measure the thickness variations in the soap films directly, we construct an interference microscope in which we illuminate the thin soap film with RGB illumination (a light source with with three narrow (~25 nm) wavelength bands around red, green and blue). We observe the colourful maps shown in Fig. 1d–f, in which the colours are true colours, exactly as the light is reflected from the thin soap film. The colours indicate the local thickness of the film (see colour map in Fig. 1c). We numerically reconstruct the thickness map (see Supplementary Information), as shown in Fig. 1g–i, and find a beautiful two-dimensional landscape of hills and valleys—a disordered but correlated thickness landscape that typically varies in the 50–550-nm range. Fig. 1g–i shows examples of different thickness landscapes, each having different correlation

---

[1]Physics Department and Solid State Institute, Technion-Israel Institute of Technology, Haifa, Israel. [2]Russell Berrie Nanotechnology Institute, Technion-Israel Institute of Technology, Haifa, Israel. [3]CREOL, The College of Optics and Photonics, University of Central Florida, Orlando, FL, USA. [4]These authors contributed equally: Anatoly Patsyk, Miguel A. Bandres. ✉e-mail: msegev@technion.ac.il; bandres@creol.ucf.edu

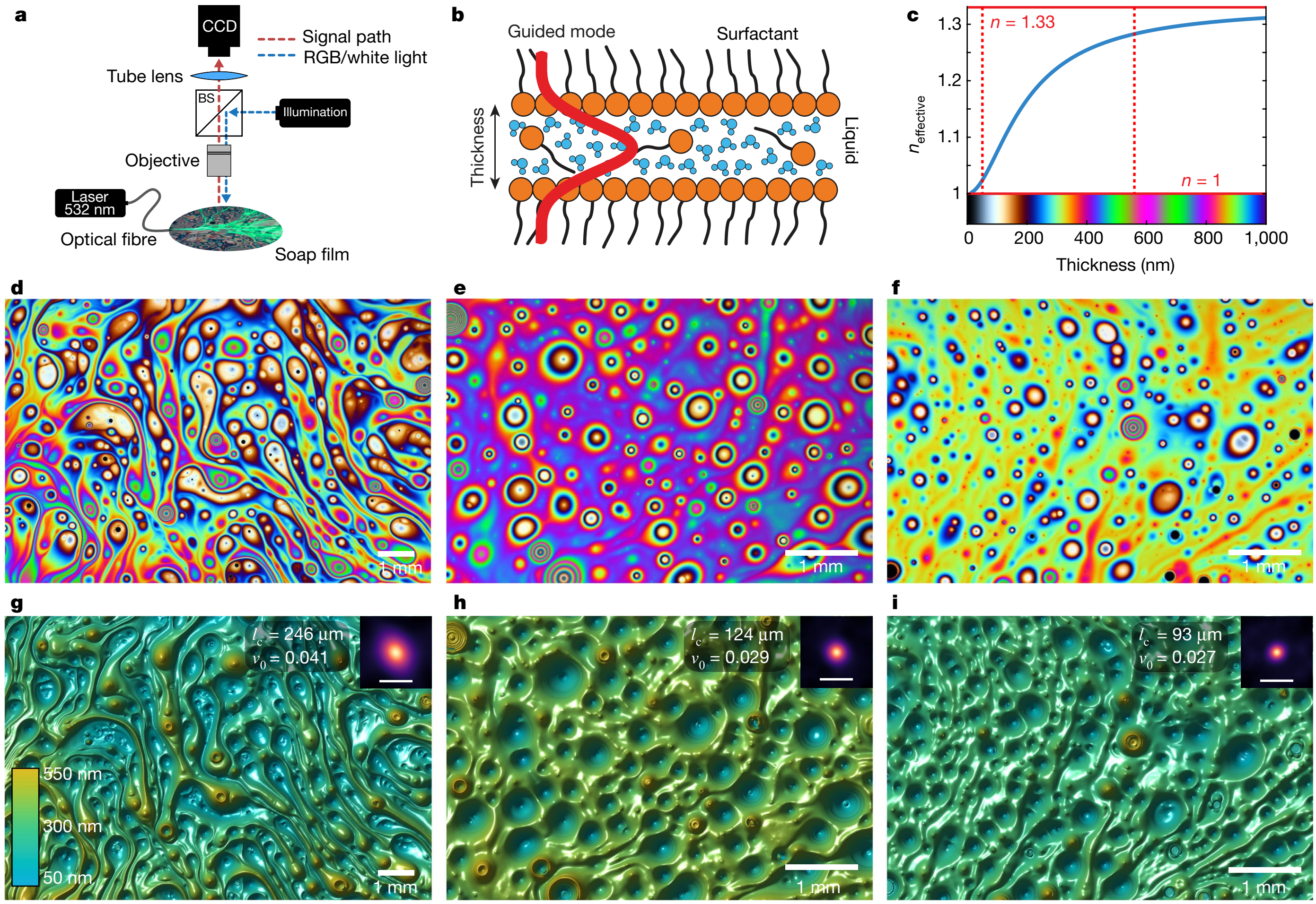


Fig. 1 | Thin liquid membranes as a platform for observing branched flow of light. a, Experimental microscope set up for observing the light propagating within a thin soap film, and the true-colour interference pattern reflected by the thin film under RGB illumination. The laser beam is coupled to the membrane by an optical fibre touching the membrane or by directly sending a collimated elliptical beam from the side of the membrane. The BS and the CCD shown in the figure refer to a beamsplitter and a charged-coupled device camera. **b**, Schematic of a thin soap film. Liquid molecules (water and/or glycerin) are held between two layers of surfactant molecules, creating a thin soap film. The film acts as a two-dimensional (slab) waveguide for the light. **c**, Effective refractive index n of the light propagating inside the film as a function of the film thickness. The red dashed lines indicate the range of thickness variation in our experiments. The colour scale shows the actual colour of the reflected RGB light at each thickness. The thickness of the membranes in the experiment is less than a micrometre. **d**–**f**, Experimental microscope images of the true-colour interference patterns created by the light reflected from the thin soap film under RGB illumination. **g**–**i**, Numerically reconstructed thickness landscape of the thin soap film from the interference colour patterns in **d**–**f**. In these three examples, the thickness variations are all in the range 50–550 nm. These thickness variations translate (through the relationship in **c**) into an effective refractive index ('potential') landscape, for the light propagating inside the thin soap film. The inset shows the autocorrelation, correlation length l_c, and strength v_0 of the effective potential. Manipulating the soap films makes it possible to produce a wide range of potential landscapes with a different l_c and v_0. The range of these parameters in our system is $v_0 \approx 1$–5% and $l_c \approx 90$–350 µm.

length and a different range of thickness variations. By manipulating the soap films—mixing or changing the surfactant/water concentration—it is possible to produce a wide range of thickness landscapes. Every membrane has a unique two-dimensional thickness landscape (a two-dimensional map). When the film is exposed to air flow in its vicinity, the thickness landscape varies over time. A membrane isolated from air movements remains stable for several minutes. The thickness landscape maps to a smooth correlated disordered effective refractive index for the light propagating within the film, through the relation in Fig. 1c. For these reasons, thin liquid films provide a perfect platform with which to observe and study the branched flow of light.

In our experimental setup, we launch a laser beam into the slab waveguide formed by a thin liquid soap film, and observe its evolution (Fig. 1a). The laser beam is coupled into the film through a single-mode fibre inserted into the film (Supplementary Fig. 2), or by coupling a broad elliptic beam (a 'plane wave' generated by a cylindrical lens) into the film. The fibre coupling is implemented by injecting the fibre into the membrane, with the fibre core aligned with the plane of the membrane slab. The fibre slightly enlarges the thickness of the membrane, but only by several micrometres in its vicinity, not affecting the rest of the membrane. The mode emitted from the fibre is much wider than the film, and hence only the first mode of the film is excited. During propagation, the beam is partially scattered from the film, which allows us to project an optical image of the light evolving in the membrane onto the camera (Fig. 1a), enabling the observation of the propagation dynamics directly in real time. As shown in Fig. 2, the beam is deflected by local random variations in the film thickness, forming focused branches that keep dividing to form a pattern that resembles the branches of a tree. The branches are created by caustics[20], which are generated when the optical wave experiences the effective refractive index landscape in the thin film. Perturbing the membrane by weak air flow in its vicinity changes the potential landscape and gives rise to different realizations of branched flow in real time, leading to the dynamic patterns shown in Supplementary Videos 1 and 2 (recorded under no illumination and

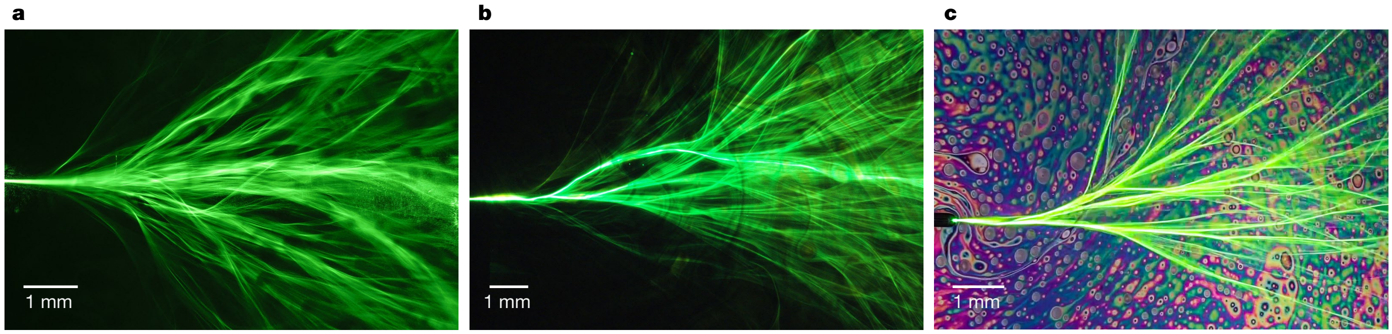


Fig. 2 | Observation of branched flow of light for an input beam generated by a single mode fibre. a, **b**, Top-view microscope images showing the evolution of a 532-nm laser beam emitted from a single mode fibre into a soap membrane. The light propagating in the film forms branched flow channelling. **c**, Branched flow pattern shown on top of the interference colour pattern generated by weak white light, making it possible to observe the potential landscape together with the branched flow.

under white light illumination using RGB sensors, respectively) for a variety of potential landscapes. By controlling the illumination intensity, we are able to observe the phenomenon of branched flow simultaneously with the underlying disordered potential landscape (Fig. 2c).

We further explore the branched flow of light under a different input beam, by launching a broad beam (approximating a plane wave) into the film and measuring its branching during propagation, as shown in Fig. 3d–f. From this figure, it is clear how the plane wave focuses at a particular distance, and how this distance varies between different landscapes of the disordered potential (Fig. 3a–c). The branched flow for a plane wave input (Fig. 3d–f) displays the expected branching of a plane wave, which was previously observed only in simulations[23–26].

Originally, branched flow was discovered in experiments with electrons travelling through a weak, smoothly varying potential in a semiconductor heterostructure, where dynamics of the electronic wavefunction was described by the time-independent Schrödinger equation

$$-\frac{\hbar^2}{2m}\nabla^2\psi + U\psi = E\psi, \qquad (1)$$

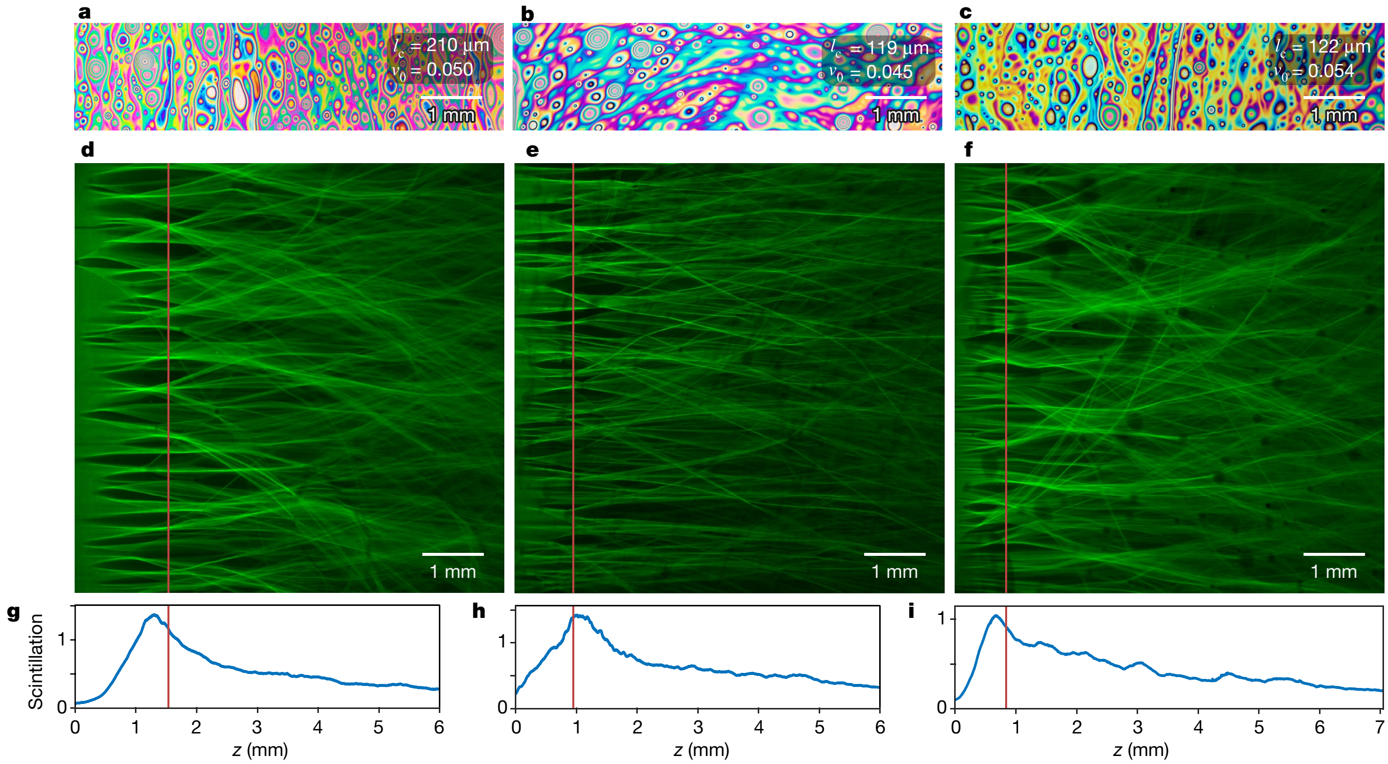


Fig. 3 | Observation of the branched flow of light for a plane wave input beam. a–**c**, Experimental microscope images of true-colour inference patterns created by the light reflected from the thin soap film under white light. The extracted values for the correlation length, l_c, and the potential strength, v_0, are given on the right. **d**–**f**, Top-view microscope images showing the amplitude (saturated at 80% of the maximum; see Supplementary Fig. 13) of the branching of a broad 532-nm laser beam as it propagates in the potential landscapes shown in **a**–**c**. **g**–**i**, Respective scintillation index, as a function of the propagation distance z, extracted from the experimental data (averaging over the transverse plane for about 10–20 realizations. The red lines in **d**–**i** mark the extracted value of $l_c v_0^{-2/3}$, which is proportional to the distance to the first branching, d_0. As shown here, we have experimentally observed that this distance, d_0, decreases as the correlation length decreases (when the potential strength is roughly the same; compare **g** and **i**). Also, d_0 decreases as the potential strength increases, (when the correlation length is roughly the same; compare **h** and **i**).

where m is the electron mass, E is the electron energy and U is the random potential in the semiconductor. In our experiments, we study the branched flow of optical waves in a slab waveguide created by a thin liquid film of thickness on the order of a single wavelength. The evolution of time-harmonic waves inside the thin film follows the Helmholtz equation

$$-\nabla^2\Psi + k_0^2(\bar{n}^2 - n_{\text{eff}}^2)\Psi = k_0^2\bar{n}^2\Psi \tag{2}$$

where $\Psi(x, z)$ is the electric field component parallel to the plane of the thin film, $k_0 = 2\pi/\lambda$ is the wavenumber in vacuum, $n_{\text{eff}}(x, z)$ is the effective refractive index for a given guided mode in the waveguide and $\bar{n}^2 = \langle n_{\text{eff}}^2\rangle$ is obtained by averaging over the whole sample; see the Supplementary Information. Equations (1) and (2) are mathematically equivalent, where for the optical wave the role of energy is played by $k_0^2\bar{n}^2$, with an effective potential $V(x, z) = k_0^2(\bar{n}^2 - n_{\text{eff}}^2)$ that has a zero mean $\langle V\rangle = 0$. As shown in the Supplementary Information, n_{eff} is a function of the local thickness of the slab, the optical wavelength and the refractive index of the film. In this way, local thickness variations in the film are manifested in a smoothly varying disordered potential landscape experienced by the optical wave propagating within the film.

Branched flow is universally characterized by two global parameters of the disordered correlated potential: the potential strength, which is the ratio between the standard deviation of the potential and the energy, $\upsilon_0 = \sqrt{\langle V^2\rangle}/2E = 0.5\sqrt{\langle n_{\text{eff}}^4\rangle/\bar{n}^4 - 1}$, and the correlation length, l_c, defined by the autocorrelation function. For our two-dimensional random potential, the autocorrelation function is $c(\mathbf{r}) = \langle V(\mathbf{r})V(0)\rangle = V_0^2 f(|\mathbf{r}|/l_c)$, with $f(0) = 1$ and $V_0 = \sqrt{\langle V^2\rangle} = \sqrt{c(0)}$. Being a universal phenomenon, branched flow is independent of the exact spatial structure of the potential and does not depend even on the form of the correlation function f, which may be any smooth function[20]. Our experimental platform of soap films allows us to generate a wide range of potential landscapes, with strengths and correlation lengths varying between $\upsilon_0 \approx 1$–5% and $l_c \approx 90$–$350\ \mu$m, respectively. Typically, these statistical parameters vary only by 5% across different sections of every film (see Supplementary Fig. 6). Examples of different thickness landscapes are shown in Fig. 1g–i, which are generated in the same system under slightly different conditions (see Supplementary Information). The statistical features of branched flow are manifested in the distance from the 'source' (input beam) to the first branching point, d_0. This distance was found[8,14,20] to satisfy $d_0 \propto l_c\upsilon_0^{-2/3}$. Since our system provides for easy generation of many realizations of the random potential, we are able to study the relation between d_0 and the parameters υ_0 and l_c using large statistical ensembles, varying the correlation length, and so on. To extract d_0, we measure the scintillation index—the normalized variance of the branched flow intensity, $S(z) = \langle I^2(z)\rangle/\langle I(z)\rangle^2 - 1$, as the branches evolve along z. Here, I is the (local) intensity and the average is taken over different realizations of random potentials with the same υ_0 and l_c (ref. [8]). The scintillation index is a convenient notion, because it peaks when fluctuations are maximal, marking the onset of branching and therefore obeying the same scaling law as the distance to the first caustic[8,20].

The measured scintillation index is shown in Fig. 3g–i, for a plane wave launched into the film. To extract the scintillation index, we average the intensity in each individual realization over the transverse coordinate, which allows convergence of $S(z)$ in just a few realizations (see Supplementary Information). As shown by Fig. 3g–i, the observed scintillation index grows sharply at a distance proportional to $l_c\upsilon_0^{-2/3}$, reaches a maximum, and then declines slowly to a constant value with a long tail. As Fig. 3 shows, the position of the peak of the scintillation index is in close proximity to the calculated value of $l_c\upsilon_0^{-2/3}$ (red line in Fig. 3d–i), thus experimentally revealing the scaling law for a variety of potential landscapes. To corroborate our experiments, we also carry out simulations, shown in the Supplementary Information, of a plane wave launched into the two-dimensional refractive index landscapes of Fig. 1g–i, constructed from the actual experimental interference colour patterns of Fig. 1d–f. The correspondence between the branched flow observed directly in the experiments (Fig. 3) and the simulated branched flow using the actual measured potential landscape (Supplementary Figs. 4, 5, 8–12) is clearly visible.

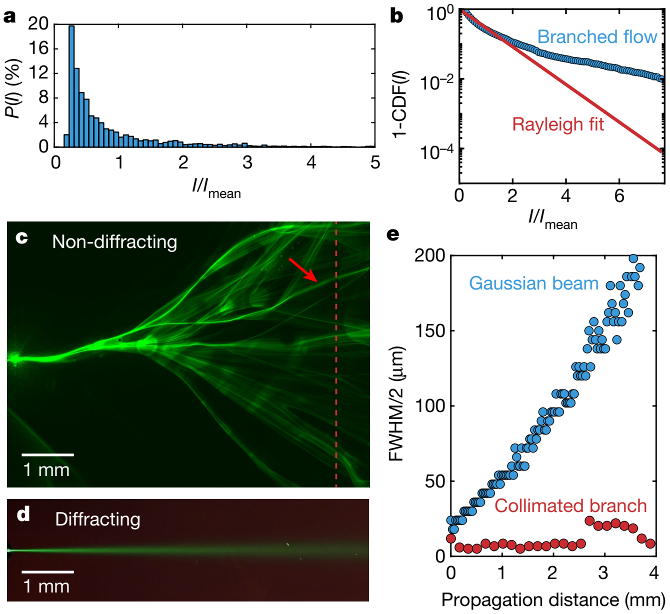


Fig. 4 | Statistical properties of experimentally observed optical branched flow for a narrow input beam. a, **b**, Probability distribution and cumulative distribution of the branches' intensities, for 100 experimental observations of branched flow, in the same film. The statistics is over the peak intensities of the branches at a fixed distance for each realization, chosen to be far enough from the input such that the branching is fully developed, as depicted in the specific realization shown in **c**, with the dashed line marking the distance from the source. As shown in **b**, the cumulative probability displays a heavy tail, as compared to that in an uncorrelated random potential (the Rayleigh fit). This implies an increased probability of finding intense waves due to the correlations in the potential landscape, as compared to an uncorrelated potential, which exhibits an exponential decay of the probability to find intense peaks. **c**, Typical experimental image of branched flow, revealing many channels where diffraction is arrested. **d**, Experimentally observed evolution of a Gaussian beam (for the same initial width as the input beam in **c**) in a flat membrane, exhibiting diffraction broadening characteristic of a homogeneous medium. **e**, Comparison between the width of the collimated branch marked by red arrow in **c** and the width of the freely diffracting Gaussian beam of **d** for a propagation interval of 4 mm.

Our experiments allow the extraction of additional statistical features of branched flow, such as the statistics of the caustic intensities. The probability density of the branched flow intensity, shown in Fig. 4a, is calculated from the imaged branched flow patterns. In this process, we measure the peak intensities, I_{peak}, along the red line in Fig. 4c, which marks a set distance from the launch point where multiple branches have already formed. We repeat this process for 100 experiments using slightly different launch positions with the same potential landscape, giving rise to 100 different branched flow patterns. We find all the intensity peaks I_{peak} at a given distance from the launch point (for example, the peaks at the plane marked by the red dashed line in Fig. 4c), identify all I_{peak} in each realization, and calculate the probability distribution (from all 100 different realizations of branched flow in the same stable membrane) shown in Fig. 4a (see details in the Supplementary Information). For a correlated potential, this statistics was predicted to display a heavy-tail distribution[9,12], whereas for a completely random potential the tail of the distribution should display exponential decay[10,14] (blue line in Fig. 4b). As exemplified by

Fig. 4b and Supplementary Fig. 7, the potential in our experiments is always correlated, and we therefore expect a considerable increase in the probability of the occurrence of localized high intensity waves. Indeed, the measured probability distribution functions displayed in Fig. 4a and b show that for low intensities, the cumulative probability follows the expected exponential decay, but that at sufficiently high intensities (above the mean intensity) the probability begins to deviate substantially from exponential decay—as the occurrence of extreme waves is increased owing to the formation of branches of high intensity by the correlated potential.

Another property of branched flow that has thus far not attracted attention is the arrest of diffraction broadening in the branches, which may be viewed as quenching of transverse diffusion. We find that the branches exhibit much less diffraction broadening than do ordinary wavepackets (beams), despite the fact that the branches are formed by scattering from random fluctuations. The branches behave as collimated narrow channels, even though the beam propagates in a random potential—in which one may expect that the beam will scatter randomly. Experimentally, it is instructive to compare the width of propagating branches from Fig. 4c to the width of the corresponding Gaussian beams propagating in a homogeneous film (Fig. 4d). Figure 4e compares the width of a characteristic branch from Fig. 4c (marked by red arrow) to the diffraction of a Gaussian beam in a film with a uniform thickness. The branched channel maintains the same width for at least ten diffraction lengths (Rayleigh lengths) before splitting again or experiencing diffraction broadening. This arrested diffraction broadening of the branches seems to be a universal feature of branched flow. Usually, nondiffracting beams are generated by nonlinear processes such as self-focusing driven by Kerr or saturable nonlinearities. Here the broadening of these wavepackets is arrested owing to scattering from the correlated random potential without nonlinear effects. Interestingly, the evolution of branched flow is fundamentally different from another phenomenon associated with random potentials: Anderson localization[27], which in the scheme of transverse localization requires the potential to be invariant in the propagation direction. Here, of course, the random variations in the potential occur anywhere in the plane, and vary in both the transverse direction and in the propagation direction. Thus, the arrest of diffraction broadening of the branches is not related to Anderson localization, despite both being generated by a random potential.

Before closing, it is important to emphasize that our experimental platform—of thin liquid films of soap—is fluidic, and so the soap film, together with the laser beam, constitutes an optofluidic system. Optofluidics, the science of light interacting with fluids, presents a host of linear and nonlinear phenomena, where light–fluid interactions give rise to effects that are fundamentally different from those encountered in light–solid interactions. The mobility of the fluid, the possibility of optically inducing deformations in the flow field, the role of diffusion and convection in transporting heat and substance, and the large-scale inhomogeneities emerging when a fluid interacts with light—all of these contribute to a variety of nonlinear phenomena that are not encountered when light interacts with solids. Examples range from nonlinearities induced by optical forces on microparticles[28], optical control of thermocapillary effects in complex nanofluids[29], particle manipulation[30] and more. In this context, using liquid soap films as a platform for experimenting with the branched flow of light has major implications for future research, such as investigating the effects of optical forces (the gradient force and radiation pressure) on branched flow. Our fluidic system may be ideal for such avenues of research, because at high enough intensities the optical forces (or heat absorption) will affect the thickness variations and perhaps create stochastic solitons. The effects of optical forces on branched flow in our thin fluidic films could offer control of flow by light and give rise to new phenomena driven by the symbiotic dynamics of the branched flow of light affecting the flow of the liquid, suggesting the occurrence of turbulence at low Reynolds numbers. Also, making the soap films slightly thicker to allow variations in the refractive index in the narrow dimension of the film, and in addition to support multiple guided modes inside the film, could give rise to branched flow in three dimensions, a phenomenon that has been proposed[31] but has thus far never been observed. In such scenarios, the full three-dimensional variation of refractive index would be required, rather than an effective refractive index.

The demonstration of the branched flow of light in our optofluidic platform of thin soap films enables access to other experimental regimes; for example, the thin soap films could be shaped into a variety of curved surfaces to study the branched flow in curved space. Supplementary Video 3 shows such an example from our experiments using a spherical shell and thus demonstrating branched flow in curved space. Such curved space experiments are intimately related to general relativity[32,33]. Moreover, when the soap film is made to be slightly absorptive, the thermal effects modify the surface tension and affect the branched flow. Likewise, if the medium displays thermal optical nonlinearity, such experiments could relate to branched flow in the Newton–Schrödinger framework of general relativity[34] in which scattering of the wavefunction has not yet been explored. Similarly, photonics offers the ability to manipulate gain and loss, and also to design parity–time-symmetric systems[35], in which branched flow has never been envisioned. Undoubtedly, the phenomenon of branched flow of light in thin liquid films suggests a plethora of ideas, and we foresee many surprising results.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2376-8.

1. Topinka, M. A. et al. Coherent branched flow in a two-dimensional electron gas. *Nature* **410**, 183–186 (2001).
2. Shaw, S. E. J. Propagation in smooth random potentials. PhD thesis, Harvard (2002); https://search.proquest.com/docview/305536276.
3. Aidala, K. E. et al. Imaging magnetic focusing of coherent electron waves. *Nat. Phys*. **3**, 464–468 (2007).
4. Jura, M. P. et al. Unexpected features of branched flow through high-mobility two-dimensional electron gases. *Nat. Phys*. **3**, 841–845 (2007).
5. Maryenko, D. et al. How branching can change the conductance of ballistic semiconductor devices. *Phys. Rev. B* **85**, 195329 (2012).
6. Liu, B. & Heller, E. J. Stability of branched flow from a quantum point contact. *Phys. Rev. Lett*. **111**, 236804 (2013).
7. Höhmann, R., Kuhl, U., Stöckmann, H.-J., Kaplan, L. & Heller, E. J. Freak waves in the linear regime: a microwave study. *Phys. Rev. Lett*. **104**, 093901 (2010).
8. Barkhofen, S., Metzger, J. J., Fleischmann, R., Kuhl, U. & Stöckmann, H.-J. Experimental observation of a fundamental length scale of waves in random media. *Phys. Rev. Lett*. **111**, 183902 (2013).
9. Heller, E. J., Kaplan, L. & Dahlen, A. Refraction of a Gaussian seaway. *J. Geophys. Res. Oceans* **113**, https://doi.org/10.1029/2008JC004748 (2008).
10. Ying, L. H., Zhuang, Z., Heller, E. J. & Kaplan, L. Linear and nonlinear rogue wave statistics in the presence of random currents. *Nonlinearity* **24**, R67 (2011).
11. Degueldre, H., Metzger, J. J., Geisel, T. & Fleischmann, R. Random focusing of tsunami waves. *Nat. Phys*. **12**, 259–262 (2016).
12. Wolfson, M. A. & Tomsovic, S. On the stability of long-range sound propagation through a structured ocean. *J. Acoust. Soc. Am*. **109**, 2693–2703 (2001).
13. Mattheakis, M., Tsironis, G. P. & Kaxiras, E. Emergence and dynamical properties of stochastic branching in the electronic flows of disordered Dirac solids. *EPL* **122**, 27003 (2018).
14. Mattheakis, M. & Tsironis, G. P. Quodons in mica. In *Extreme Waves and Branched Flows in Optical Media* 425–454 (Springer International Publishing, 2015).
15. Mattheakis, M., Pitsios, I. J., Tsironis, G. P. & Tzortzakis, S. Extreme events in complex linear and nonlinear photonic media. *Chaos Solitons Fractals* **84**, 73–80 (2016).
16. Dudley, J. M., Dias, F., Erkintalo, M. & Genty, G. Instabilities, breathers and rogue waves in optics. *Nat. Photon*. **8**, 755–764 (2014).
17. Akhmediev, N., Soto-Crespo, J. M. & Ankiewicz, A. Extreme waves that appear from nowhere: on the nature of rogue waves. *Phys. Lett. A* **373**, 2137–2145 (2009).
18. Brandstötter, A., Girschik, A., Ambichl, P. & Rotter, S. Shaping the branched flow of light through disordered media. *Proc. Natl Acad. Sci. USA* **116**, 13260–13265 (2019).
19. Berry, M. V. & Upstill, C. in *Progress in Optics. IV. Catastrophe Optics: Morphologies of Caustics and Their Diffraction Patterns* (ed. Wolf, E.) Vol. 18, 257–346 (Elsevier, 1980).

20. Kaplan, L. Statistics of branched flow in a weak correlated random potential. *Phys. Rev. Lett*. **89**, 184103 (2002).
21. Kravtsov, Y. A. & Orlov, Y. I. *Caustics, Catastrophes and Wave Fields* (Springer, 1993).
22. Patsyk, A., Bandres, M. A. & Segev, M. Interaction of light with thin liquid membranes. In *Conference on Lasers and Electro-Optics* FF3E2 (Optical Society of America, 2018); https://www.osapublishing.org/abstract.cfm?uri=CLEO_QELS-2019-FTu3D.1.
23. Metzger, J. J., Fleischmann, R. & Geisel, T. Universal statistics of branched flows. *Phys. Rev. Lett*. **105**, 020601 (2010).
24. Degueldre, H., Metzger, J. J., Schultheis, E. & Fleischmann, R. Channeling of branched flow in weakly scattering anisotropic media. *Phys. Rev. Lett*. **118**, 024301 (2017).
25. Metzger, J. J., Fleischmann, R. & Geisel, T. Statistics of extreme waves in random media. *Phys. Rev. Lett*. **112**, 203903 (2014).
26. Metzger, J. J., Fleischmann, R. & Geisel, T. Intensity fluctuations of waves in random media: what is the semiclassical limit? *Phys. Rev. Lett*. **111**, 013901 (2013).
27. Schwartz, T., Bartal, G., Fishman, S. & Segev, M. Transport and Anderson localization in disordered two-dimensional photonic lattices. *Nature* **446**, 52–55 (2007).
28. Ashkin, A., Dziedzic, J. M. & Smith, P. W. Continuous-wave self-focusing and self-trapping of light in artificial Kerr media. *Opt. Lett*. **7**, 276–278 (1982).
29. Lamhot, Y. et al. Optical control of thermocapillary effects in complex nanofluids. *Phys. Rev. Lett*. **103**, 264503 (2009).
30. Schley, R. et al. Loss-proof self-accelerating beams and their use in non-paraxial manipulation of particles' trajectories. *Nat. Commun*. **5**, 5189 (2014).
31. Heller, E. J., Fleischmann, R. & Kramer, T. Branched flow. Preprint at https://arxiv.org/abs/1910.07086 (2019).
32. Bekenstein, R., Nemirovsky, J., Kaminer, I. & Segev, M. Shape-preserving accelerating electromagnetic wave packets in curved space. *Phys. Rev. X* **4**, 011038 (2014).
33. Patsyk, A., Bandres, M. A., Bekenstein, R. & Segev, M. Observation of accelerating wave packets in curved space. *Phys. Rev. X* **8**, 011001 (2018).
34. Bekenstein, R., Schley, R., Mutzafi, M., Rotschild, C. & Segev, M. Optical simulations of gravitational effects in the Newton–Schrödinger system. *Nat. Phys*. **11**, 872–878 (2015).
35. El-Ganainy, R., Makris, K. G., Christodoulides, D. N. & Musslimani, Z. H. Theory of coupled optical PT-symmetric structures. *Opt. Lett*. **32**, 2632–2634 (2007).

Data availability

The data are available upon request.

Acknowledgements We are indebted to M. V. Berry, for suggesting to us that what we saw in the experiments is related to branched flow. This research was supported by the German-Israeli DIP project, and by the Israel Science Foundation.

Author contributions All authors contributed substantially to this work.

Competing interests The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2376-8.
Correspondence and requests for materials should be addressed to M.S. or M.A.B.
Reprints and permissions information is available at http://www.nature.com/reprints.

Chemical gradients in human enamel crystallites

https://doi.org/10.1038/s41586-020-2433-3

Received: 23 August 2019

Accepted: 8 April 2020

Karen A. DeRocher[1,7], Paul J. M. Smeets[1,7], Berit H. Goodge[2,3], Michael J. Zachman[2,6], Prasanna V. Balachandran[4,5], Linus Stegbauer[1], Michael J. Cohen[1], Lyle M. Gordon[1], James M. Rondinelli[1], Lena F. Kourkoutis[2,3] & Derk Joester[1]✉

Dental enamel is a principal component of teeth[1], and has evolved to bear large chewing forces, resist mechanical fatigue and withstand wear over decades[2]. Functional impairment and loss of dental enamel, caused by developmental defects or tooth decay (caries), affect health and quality of life, with associated costs to society[3]. Although the past decade has seen progress in our understanding of enamel formation (amelogenesis) and the functional properties of mature enamel, attempts to repair lesions in this material or to synthesize it in vitro have had limited success[4–6]. This is partly due to the highly hierarchical structure of enamel and additional complexities arising from chemical gradients[7–9]. Here we show, using atomic-scale quantitative imaging and correlative spectroscopies, that the nanoscale crystallites of hydroxylapatite ($Ca_5(PO_4)_3(OH)$), which are the fundamental building blocks of enamel, comprise two nanometric layers enriched in magnesium flanking a core rich in sodium, fluoride and carbonate ions; this sandwich core is surrounded by a shell with lower concentration of substitutional defects. A mechanical model based on density functional theory calculations and X-ray diffraction data predicts that residual stresses arise because of the chemical gradients, in agreement with preferential dissolution of the crystallite core in acidic media. Furthermore, stresses may affect the mechanical resilience of enamel. The two additional layers of hierarchy suggest a possible new model for biological control over crystal growth during amelogenesis, and hint at implications for the preservation of biomarkers during tooth development.

Enamel covers the entire crown of human teeth (Fig. 1a), reaching thicknesses of several millimetres (Fig. 1b). A characteristic microstructural element, the enamel rod (Fig. 1c), is composed of thousands of lath-like crystallites aligned with their crystallographic *c* direction approximately parallel to the long axes of the rods (Fig. 1d). Crystallites sectioned normal to their long axis appear as oblong polygons with an edge length of 20–50 nm in the short direction and 70–170 nm in the long direction (Fig. 1e, f). Characteristic length scales of the periodic hydroxylapatite (OHAp) lattice are in the subnanometre regime (Fig. 1g–i).

Enamel owes its hardness (up to about 5 GPa) to its high mineral content (approximately 96 wt%)[10]. Although enamel is nominally composed of OHAp, magnesium (0.2–0.6 wt%), sodium (0.2–0.9 wt%), carbonate (2.7–5 wt%) and fluoride (about 0.01 wt%) are also present[11]. Although the distribution of the minor constituents is known to vary over tens to hundreds of micrometres[11], gradients over much shorter distances have only recently been discovered. Specifically, in rodent incisor enamel, most Mg is confined between crystallites as Mg-substituted amorphous calcium phosphate (Mg-ACP), controlling enamel dissolution and mechanical properties[7]. Segregation of Mg and Na ions to a 2–10-nm-thick layer between human enamel crystallites was confirmed[8]. However, it has not yet been shown that this layer is identical to the Mg-ACP found in rodent enamel.

Perhaps uniquely to human enamel, the centre of the crystallite seems to be more soluble, is more prone to electron-beam-induced damage, and displays a poorly understood contrast feature that is known as the central dark line (CDL)[12,13]. All three have generally been assumed to be related to the presence of defects in the crystallite lattice, but the exact nature of these defects is not known[13]. We therefore set out to test whether there are compositional gradients of minor enamel constituents across single crystallites.

In annular dark-field scanning transmission electron microscopy (STEM-ADF) images of human outer enamel, crystallites are separated by narrow regions that appear darker than the crystallite (Fig. 1e, f), consistent with expectations for a Mg-rich amorphous intergranular phase. Additionally, they exhibit a shell that appears brighter than the core (Fig. 1e, f). The sensitivity of enamel to radiation limits the tolerable electron dose and hampers high-resolution analysis of

---

[1]Department of Materials Science and Engineering, Northwestern University, Evanston, IL, USA. [2]School of Applied and Engineering Physics, Cornell University, Ithaca, NY, USA. [3]Kavli Institute at Cornell for Nanoscale Science, Cornell University, Ithaca, NY, USA. [4]Department of Materials Science and Engineering, University of Virginia, Charlottesville, VA, USA. [5]Department of Mechanical and Aerospace Engineering, University of Virginia, Charlottesville, VA, USA. [6]Present address: Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, TN, USA. [7]These authors contributed equally: Karen A. DeRocher, Paul J. M. Smeets. ✉e-mail: d-joester@northwestern.edu

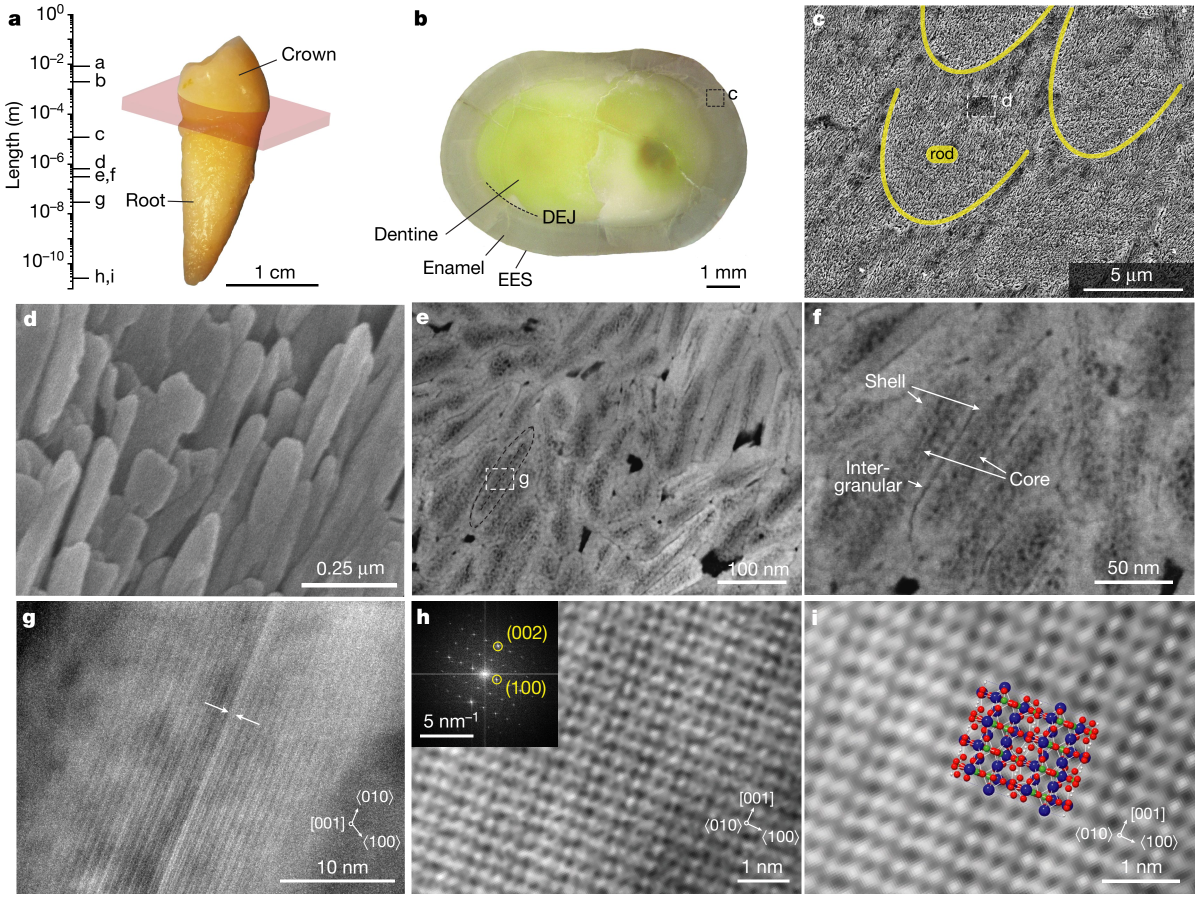


Fig. 1 | The hierarchical architecture of human enamel. a, Left, length scales of enamel in a human premolar (values for **a**–**i** are indicated); right, optical image. **b**, Section parallel to the mid-coronal cervical plane (indicated in pink in **a**) with the external enamel surface (EES) and the dentine-enamel junction (DEJ) labelled. **c**, SEM image of keyhole-shaped cross-sections of enamel rods in lactic-acid-etched outer enamel; location of image shown boxed in **b**. **d**, SEM image of OHAp crystallites. **e**–**g**, STEM-ADF images at increasing magnification of enamel crystallites in cross-section, oriented approximately parallel to the [001] zone axis (shown in **g**). Arrows in **f** indicate the intergranular space between crystallites, and shell and core regions of crystallites. The CDL in **g** (arrowed) appears bright in ADF. **h**, Cryo-STEM-ADF lattice image of a crystallite oriented parallel to the [010] zone axis (inset, fast Fourier transform). **i**, Close-up of **h** with rendering of a 2 × 2 × 2 OHAp supercell (Ca, blue; O, red; P, green; H, white).

enamel crystallites (Supplementary Information section 1.3). Here, sample cooling and low-dose imaging conditions in cryogenic STEM (cryo-STEM) enable atomic-resolution imaging of ultrathin sections (20–30 nm; Supplementary Fig. 1, Supplementary Table 1), revealing a continuous atomic lattice across the entire crystallite (Figs. 1h, i, 2a; Supplementary Figs. 2, 3). Unlike the shell, the core appears as a patchwork of lighter and darker areas on either side of the CDL, and seems more prone to beam damage[13].

Although STEM in conjunction with energy-dispersive X-ray spectroscopy (STEM-EDS) reported approximately 0.4 atom % (at.%) Mg, 0.7 at.% Na and 0.6 at.% F in crystallites (Supplementary Fig. 4, Supplementary Table 2), elemental maps were largely featureless (not shown). The presence of at least some Mg on apatite sites in addition to Mg-ACP was confirmed by bulk X-ray absorption spectroscopy (XAS) at the Mg K-edge (Fig. 2b, Supplementary Fig. 7). Major components are well represented in cryo-STEM electron energy-loss spectra (EELS; Fig. 2c–e). Unfortunately, Na and F were not detected, the C K-edge is rather weak (Fig. 2e inset), and spectral component maps did not show gradients within crystallites (Supplementary Fig. 5), despite the use of a sensitive direct electron detector for EELS[14].

However, a feature near the expected onset of the Mg $L_{2,3}$-edge at 51 eV decomposes into two principal components (with no residual) by multivariate curve resolution (MCR, Fig. 2f, Supplementary Fig. 6)[15]. One of these is predominant in the shell, the other in the intergranular Mg-ACP and in the core (compare Fig. 2g, h, Supplementary Fig. 6). Other elements may also have minor edges in this spectral area, whereas STEM-EDS and bulk XAS do support the presence of Mg. The two components suggest distinct local environments in the core and shell, and even gradients within the core (Fig. 2i). As susceptibility to beam damage limited our ability to further explore such gradients by cryo-STEM methods, we turned to atom probe tomography (APT, Supplementary Fig. 8)[16,17] to provide additional insights.

APT spectra of untreated human enamel resemble those of rodent enamel (Supplementary Fig. 9, Supplementary Table 4)[7,16]. In 3D reconstructions of APT data, it takes some practice to recognize the faceted cross-sections of enamel crystallites (Supplementary Figs. 10, 11). Treatment of samples with 250 mM NaF at pH 8.4 (37 °C, 24 h), however, resulted in a marked increase of CaF^+ ions (Supplementary Table 5). These ions clearly outline individual crystallites (compare Supplementary Fig. 10a–c, and d, e) and thus greatly facilitate the analysis of reconstructions (Fig. 3). Fluoride treatment further results in an increase in sodium content (+0.39 at.%, +71% relative) with moderate statistical significance ($P < 0.1$) that also seems to be limited to the intergranular phase (see below), but has a negligible overall impact on composition of the samples otherwise (Supplementary Tables 5, 6). These observations confirm that Na^+ and F^- rapidly diffuse between, but do not appreciably penetrate into, crystallites under the treatment conditions. Similar to rodent enamel, Mg is enriched in this intergranular

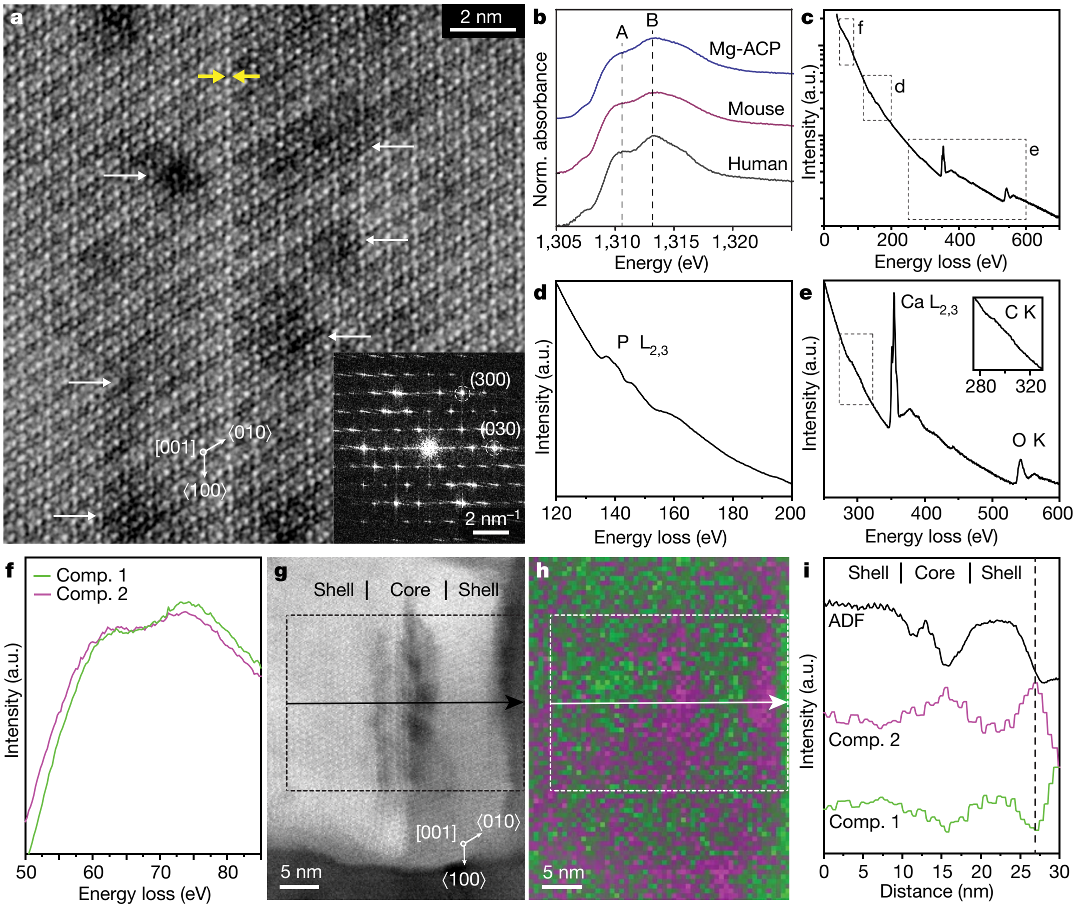


Fig. 2 | Atomic-scale structure and composition of human enamel crystallites. a, Spherical-aberration (Cs)-corrected cryo-STEM-ADF lattice image of the core of a single enamel crystallite oriented to the [001] zone axis, close to the CDL (yellow arrows). White arrows indicate darker patches in the core region. Inset, fast Fourier transform. **b**, Mg K-edge XANES of human enamel and reference materials. Peaks at approximately 1,311 eV (A) and 1,313 eV (B) are more pronounced in spectra of human enamel and indicate the presence of Mg on OHAp lattice sites. Fit parameters are reported in Supplementary Table 3. **c**–**e**, Cryo-STEM EEL spectra obtained from a region containing several enamel crystallites, with close-ups of the P $L_{2,3}$-edge (**d**) and the Ca $L_{2,3}$, O K and C K-edges (**e**). **f**, MCR components (Comp. 1,2) contributing to feature near the Mg $L_{2,3}$-edge. **g**, Cryo-STEM-ADF image of an enamel crystallite. **h**, Spatial intensity map of MCR components 1 (green) and 2 (magenta) in **g**. **i**, Average intensity profile for the region of interest indicated in **g** and **h**, in the direction of the arrow.

space (mean 0.35 at.%; range 0.15–0.51 at.%); in fluoridated samples levels of Na (1.27 at.%, 0.69–1.76 at.%) and F (1.36 at.%, 1.10–1.59 at.%) are also elevated there (compare Fig. 3e versus f; see also Supplementary Figs. 12, 13). In combination with the disordered local structure around Mg observed by Mg K-edge XAS, this is robust evidence for the presence of Mg-ACP in the intergranular phase; the thickness of this region is consistent with previous observations in human enamel[8].

In striking difference to rodent enamel, however, Mg levels are elevated not only in the intergranular Mg-ACP, but also in two distinct layers in the core (Fig. 3a; Supplementary Figs. 10, 11; Supplementary Videos 1, 2). The core is further enriched in sodium, probably as Na^+, and fluorine, probably as F^- (Fig. 3b, c). In addition, the carbon concentration is elevated (Fig. 3d), which is most probably due to the presence of carbonate (CO_3^{2-}). APT data thus prove true the hypothesis that there exist a Mg^{2+}- and CO_3^{2-}-rich core in human enamel crystallites[18,19].

Line profiles taken approximately normal to the midplane in 20 crystallites identified in 3D reconstructions of five APT sample tips (three NaF-treated, two untreated; Fig. 3e, f, Supplementary Figs. 12, 13) reveal that, on average, the two Mg-rich layers (mean 0.5 at.%; range 0.33–0.72 at.%) are also enriched in Na. However, Na levels usually peak closer to the midplane (1.2 at.%; 0.87–1.55 at.%), where F (1.4 at.%; 1.13–2.44 at.%) and C (about 0.6 at.%; 0.45–1.01 at.%) are also elevated, and Mg goes through a minimum (0.4 at.%). We note that the distributions of Na, F and C are more variable than that of Mg and can be asymmetric or show additional local maxima. Although a form of shot noise may be responsible for the latter effect, we note that contrast in STEM-ADF shows similar variation. It is therefore possible that there is some clustering of substituents. Regardless, mole fractions are always noticeably lower in the shell (on average, 0.22 at.% Mg, 0.81 at.% F, 0.52 at.% Na and 0.32 at.% C). APT therefore not only confirms the core–shell structure observed in many crystallites and over large areas by STEM, but clearly indicates that the core itself has a sandwich structure.

Driven by the size mismatch between Ca^{2+} and Mg^{2+}, we expect a substantial contraction of the apatite lattice in the Mg-rich layers[20–23]. Lattice parameters determined by density functional theory (DFT) calculations and X-ray diffraction experiments, after correction for thermal expansion, agree within 1% and indicate a contraction in both the *a* and *c* directions with increasing level of Mg substitution (Supplementary Tables 7, 8; Supplementary Figs. 14–16)[24]. Carbonate substitution also results in a contraction of the lattice in the *a* directions[25]. However, there is a mild expansion of the lattice in the *c* direction that partially offsets the effect of Mg (Supplementary Fig. 17).

Because enamel crystallites are coherent, lattice parameter changes that result from gradients in Mg^{2+} and CO_3^{2-} are in effect residual (eigen) strains and may therefore cause a net residual stress. Residual stresses in turn can affect the overall mechanical performance of a material,

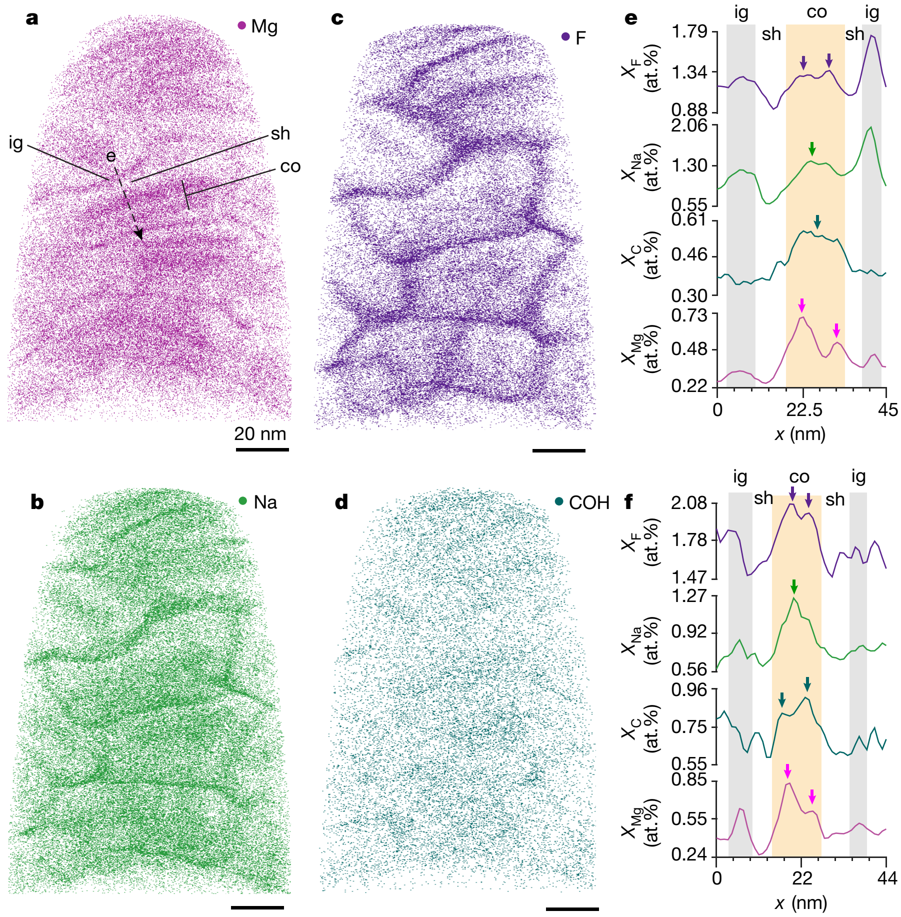


Fig. 3 | Chemical gradients in human enamel crystallites and the amorphous intergranular phase. a–**d**, Rendering of Mg (**a**), Na (**b**), F (**c**) and COH (**d**) positions in a 3D reconstruction of fluoridated human enamel, viewed along the long axis of the crystallites. All scale bars, 20 nm. **e**, Concentration profiles of F (purple), Na (green), C (teal) and Mg (magenta) along the dashed line in **a**. Profiles for *n* = 15 crystallites across 3 technical replicates are shown in Supplementary Fig. 13a–o. **f**, Concentration profiles in a crystallite from a sample that had not been fluoridated. Profiles for *n* = 5 crystallites across 2 biological replicates are shown in Supplementary Fig. 13p–t. Note that fluoridation increases the concentration of Na and F in the intergranular phase (ig, grey highlights) versus the core (co, orange highlight), due to short circuit diffusion, whereas the concentration in the shell (sh) is not affected[7]. Coloured arrows indicate local maxima of concentrations in the core.

but also affect the local chemical potential and therefore the solubility. To explore these possibilities, we predicted residual stresses in an idealized crystallite (Supplementary Fig. 18) using finite-element modelling (Fig. 4a–c). In mechanical equilibrium, the core experiences a net tensile stress, with distinct maxima in the Mg-rich layers (point T in Fig. 4a). The highest compressive stress (−46.4 MPa) is found on the free surface parallel to the (001) plane (Fig. 4b), again in the Mg-rich layers. The shell of the crystallite experiences compressive residual stresses (Fig. 4c). On most of the surface, these stresses are near −39 MPa. Although absolute values of the stresses reported here will vary as real crystallites differ in shape and composition from the highly idealized model we use here, we believe that the model captures trends quite well and provides insights into how we can expect crystallites to behave on average.

For instance, at the water-accessible endcap (Fig. 4b), stresses in the core are between 4 and 40 times higher than in the shell. This is expected to increase the solubility of the core compared to the shell. Indeed, the core of outer enamel crystallites is preferentially etched, similar to the intergranular Mg-ACP (Fig. 4d, Supplementary Fig. 19), and consistent with reports for crystallites extracted from human caries lesions[26,27]. The core–shell architecture and associated residual stresses are thus an important aspect of disease progression, and may be used for the modelling of dissolution and re-precipitation during the progression of caries lesions. Compressive stresses in the shell may further impede crack initiation, extend the size range at which crystallites perform at their theoretical strength[28], and deflect cracks, thereby increasing the tolerance of enamel[29,30]. On a different note, the strong modulation of stresses and the resulting strain just beneath the surface of the endcaps may be responsible for the CDL feature observed in electron microscopy. We note that although the Mg-rich layers do appear to be parallel to the CDL, proof that the CDL runs between the layers would require correlative imaging of the same crystallite by STEM and APT. Although correlative imaging of this kind is not unprecedented, it is a substantial challenge, especially for beam-sensitive materials such as enamel[31,32]. It would, however, be particularly rewarding, as one could also address the spatial correlation between contrast in STEM-ADF and local concentration maxima reported by APT.

Taken together, we find strong evidence that the core–shell architecture and resulting residual stresses affect the dissolution behaviour of human enamel crystallites and provide a plausible avenue for extrinsic toughening of enamel. This leads us to the question of how the gradients are created in the first place. During amelogenesis, mineral first precipitates in the organic enamel matrix as ribbons of amorphous calcium phosphate (ACP)[33–35]. ACP is tolerant of impurities, and it is conceivable that the ribbons retain Na^+ and CO_3^{2-} as they crystallize by an unknown mechanism. Crystallites initially grow much more slowly in thickness than in width. In human primary teeth for instance, ribbons are 3 nm thick and 29 nm wide at a distance of 25 µm from the ameloblasts, and

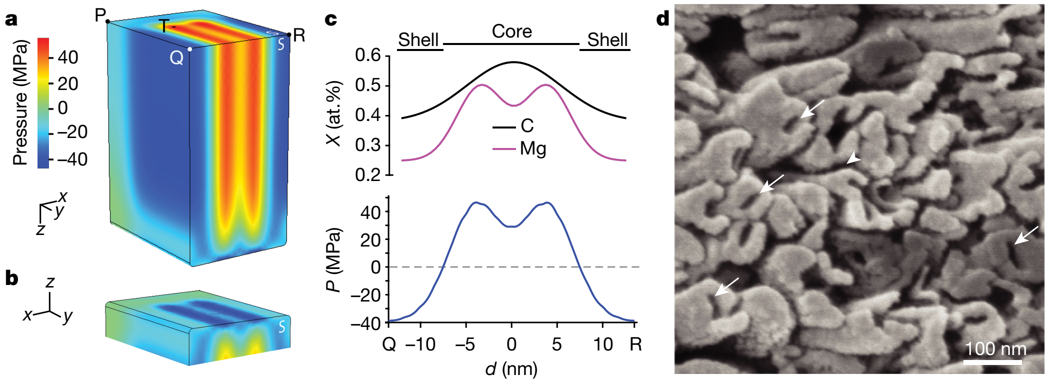


Fig. 4 | Effect of substitution on mechanical and chemical properties of human enamel crystallites. a, Rendering of the scalar pressure, calculated as one-third of the trace of the stress tensor, as a measure of residual stress in a finite-element model of an enamel crystallite. Note that symmetric boundary conditions were applied to two faces (white 'S'); values on these represent internal rather than surface stresses. **b**, View of **a** showing the free surface parallel to the (001) plane. **c**, Plot of the mole fractions of C (black) and Mg (magenta), and of the residual pressure (blue), against the distance from Q to R in **a**. **d**, SEM image of an acid-etched enamel section in which crystallites emerge end-on, displaying intergranular corrosion (arrowhead) and preferential dissolution of the core (white arrows).

grow to 10 nm thick and 58 nm wide by the time the ameloblasts have moved an additional 175 µm (Supplementary Fig. 20)[33].

In this first phase, the fast growth direction is thus parallel to the Mg-rich layers (Fig. 5; for relative growth velocities, see Supplementary Table 9), suggesting that Mg substitution breaks the crystal symmetry. Presumably, Mg acts on its own, or in concert with organic matrix molecules, by blocking active sites for growth in the direction normal to the layer direction, through anisotropic stresses generated as it is incorporated into the crystal, or a combination of these effects. This would require Mg to be deliberately introduced into the system after the ribbon has formed, and indeed the Mg concentration in porcine enamel is known to reach a maximum in the late secretory stage[36]. Consistent with a regulatory role, perturbation of putative Mg transporters is known to affect amelogenesis, even though the impact on crystallite shape is not known[37–39].

In the second phase, which is probably identical to the maturation stage of amelogenesis, growth slows down. At the same time, the ratio of the growth velocities changes, with crystals growing thicker rather than wider. In human primary enamel, this results in mature crystallites with a thickness of 26 nm and a width of 80 nm (ref. [33]). The Mg-poor shell is probably formed during this period. Slow growth at low supersaturation, which may be combined with a drop in Mg and Na concentration in the enamel matrix[36,40], is indeed expected to result in low rates of incorporation of impurities on the apatite lattice. Any Mg still present would accumulate ahead of the interface of the growing crystallite.

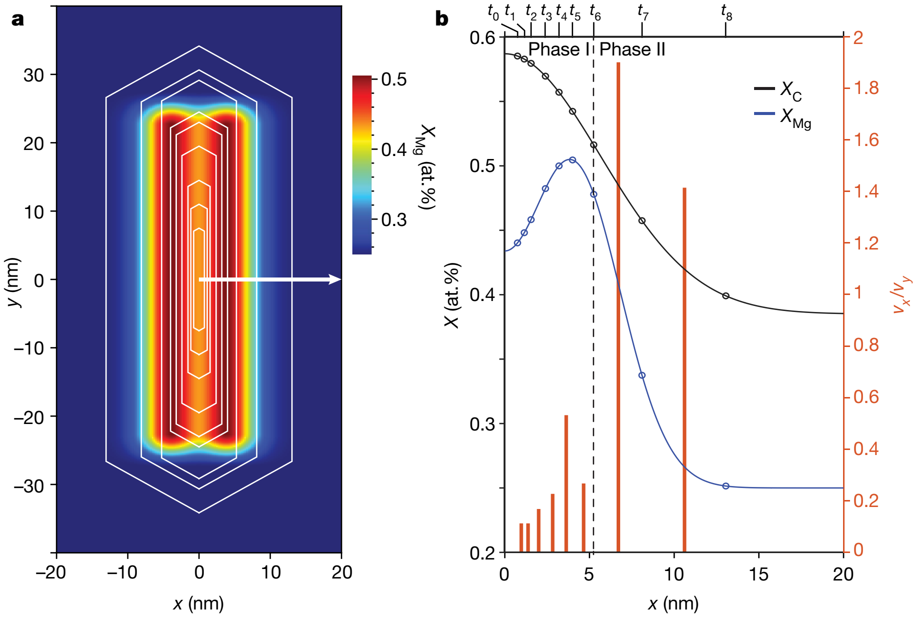


Fig. 5 | A model for human enamel crystallite growth during amelogenesis. a, Schematic drawing of growth stages (time points t_0–t_8, shown in **b**) of human primary enamel crystallites (white hexagons, after ref. [33]) superimposed on an idealized map of the Mg concentration based on observation of human permanent enamel crystallites reported herein. **b**, On the left *y* axis, plot of the mole fraction of Mg (blue) and carbon (present as carbonate; black, see Supplementary Fig. 21 for map) against distance along the white arrow in **a**. The open circles indicate the mole fractions at the interface of the growing crystallite at t_0–t_8. On the right *y* axis, plot of the ratio of the average growth velocities in the *x* and *y* directions in successive time intervals (Supplementary Table 9). Note that scaling of the time axis is unknown, and probably nonlinear. As a consequence, absolute speeds cannot be determined and may vary.

At the end of the maturation stage, as crystals start to impinge on one another, the Mg/Ca ratio in the matrix would rise rapidly, and when the threshold for Mg-ACP is exceeded[23], precipitation of the amorphous intergranular phase is triggered. Although this mechanism is consistent with and provides an explanation for a number of independent observations, we note that the data we draw on come from both primary and permanent human enamel and include observations in other species. Clearly, a thorough compositional analysis of immature enamel ribbons and crystallites in one species would greatly aid in confirming the proposed sequence of events.

Regardless of whether this mechanistic proposal is accurate, what emerges is that the concentration of Mg and other minor ions at the surface of the crystallite, and therefore also the medium in its immediate proximity, varies systematically during amelogenesis of human permanent teeth. This may affect how enamel matrix proteins and their degradation products, thought to be involved in controlling enamel crystallite formation, interact with the mineral phase and each other.

The fact that ions not essential for amelogenesis, such as fluoride, are incorporated into the crystallite core has an important corollary. Enamel forms over very specific times during the development of different teeth (in humans it starts as early as the second trimester in utero and continues until the late teens), is not appreciably remodelled, and is very well preserved in remains and fossils. Crystallite cores might thus encapsulate spatially resolved biomarkers for environmental exposure, disease, or medical intervention, over an extended period of time. With APT and correlative imaging and spectroscopy, this record is now accessible and might help decipher for instance genetic predisposition to caries or the mechanism behind molar-incisor hypomineralization (MIH), a dental developmental defect of unclear aetiology that affects as many as 20% of all schoolchildren[2,41].

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2433-3.

1. Nanci, A. *Ten Cate's Oral Histology: Development, Structure, and Function* 7th edn (Mosby Elsevier, 2008).
2. Chai, H., Lee, J. J. W., Constantino, P. J., Lucas, P. W. & Lawn, B. R. Remarkable resilience of teeth. *Proc. Natl Acad. Sci. USA* **106**, 7289–7293 (2009).
3. Klein, O. D. et al. Meeting report: a hard look at the state of enamel research. *Int. J. Oral Sci*. **9**, e3 (2017).
4. Moradian-Oldak, J. Protein-mediated enamel mineralization. *Front. Biosci*. **17**, 1996–2023 (2012).
5. Lacruz, R. S., Habelitz, S., Wright, J. T. & Paine, M. L. Dental enamel formation and implications for oral health and disease. *Physiol. Rev*. **97**, 939–993 (2017).
6. Robinson, C. Enamel maturation: a brief background with implications for some enamel dysplasias. *Front. Physiol*. **5**, 388 (2014).
7. Gordon, L. M. et al. Amorphous intergranular phases control the properties of rodent tooth enamel. *Science* **347**, 746–750 (2015).
8. La Fontaine, A. et al. Atomic-scale compositional mapping reveals Mg-rich amorphous calcium phosphate in human dental enamel. *Sci. Adv*. **2**, e1601145 (2016).
9. Gordon, L. M. & Joester, D. Mapping residual organics and carbonate at grain boundaries and in the amorphous interphase in mouse incisor enamel. *Front. Physiol*. **6**, 57 (2015).
10. Zhang, Y.-R., Du, W., Zhou, X.-D. & Yu, H.-Y. Review of research on the mechanical properties of the human tooth. *Int. J. Oral Sci*. **6**, 61–69 (2014).
11. Robinson, C. et al. The chemistry of enamel caries. *Crit. Rev. Oral Biol. Med*. **11**, 481–495 (2000).
12. Yanagisawa, T. & Miake, Y. High-resolution electron microscopy of enamel-crystal demineralization and remineralization in carious lesions. *J. Electron Microsc*. **52**, 605–613 (2003).
13. Reyes-Gasga, J., Hemmerle, J. & Bres, E. F. Aberration-corrected transmission electron microscopy study of the central dark line defect in human tooth enamel crystals. *Microsc. Microanal*. **22**, 1047–1055 (2016).
14. Hart, J. L. et al. Direct detection electron energy-loss spectroscopy: a method to push the limits of resolution and sensitivity. *Sci. Rep*. **7**, 8243 (2017).
15. de Juan, A. & Tauler, R. Multivariate curve resolution (MCR) from 2000: progress in concepts and applications. *Crit. Rev. Anal. Chem*. **36**, 163–176 (2006).
16. Gordon, L. M., Tran, L. & Joester, D. Atom probe tomography of apatites and bone-type mineralized tissues. *ACS Nano* **6**, 10667–10675 (2012).
17. Gault, B., Moody, M. P., Cairney, J. M. & Ringer, S. P. *Atom Probe Microscopy* (Springer, 2012).
18. Robinson, C., Kirkham, J., Brookes, S. J., Bonass, W. A. & Shore, R. C. *Int. J. Dev. Biol*. The chemistry of enamel development. **39**, 145–152 (1995).
19. Robinson, C., Weatherell, J. A. & Hallsworth, A. S. Distribution of magnesium in mature human enamel. *Caries Res*. **15**, 70–77 (1981).
20. Laurencin, D. et al. Magnesium incorporation into hydroxyapatite. *Biomaterials* **32**, 1826–1837 (2011).
21. Shannon, R. D. & Prewitt, C. T. Effective ionic radii in oxides and fluorides. *Acta Crystallogr. B* **25**, 925–946 (1969).
22. Hughes, J. M., Cameron, M. & Crowley, K. D. Structural variations in natural F, OH, and Cl apatites. *Am. Mineral*. **74**, 870–876 (1989).
23. LeGeros, R. Z., Sakae, T., Bautista, C., Retino, M. & LeGeros, J. P. Magnesium and carbonate in enamel and synthetic apatites. *Adv. Dent. Res*. **10**, 225–231 (1996).
24. Ben Abdelkader, S., Khattech, I., Rey, C. & Jemal, M. Synthése, caractérisation et thermochimie d'apatites calco-magnésiennes hydroxylées et fluorées. *Thermochim. Acta* **376**, 25–36 (2001).
25. Deymier, A. C. et al. Protein-free formation of bone-like apatite: new insights into the key role of carbonation. *Biomaterials* **127**, 75–88 (2017).
26. Voegel, J. C. & Frank, R. M. Stages in dissolution of human enamel crystals in dental caries. *Calcif. Tissue Res*. **24**, 19–27 (1977).
27. Tohda, H., Takuma, S. & Tanaka, N. Intercrystalline structure of enamel crystals affected by caries. *J. Dent. Res*. **66**, 1647–1653 (1987).
28. Gao, H. J., Ji, B. H., Jager, I. L., Arzt, E. & Fratzl, P. Materials become insensitive to flaws at nanoscale: lessons from nature. *Proc. Natl Acad. Sci. USA* **100**, 5597–5600 (2003).
29. Yahyazadehfar, M. et al. On the mechanics of fatigue and fracture in teeth. *Appl. Mech. Rev*. **66**, 030803 (2014).
30. Yilmaz, E. D., Schneider, G. A. & Swain, M. V. Influence of structural hierarchy on the fracture behaviour of tooth enamel. *Phil. Trans. R. Soc. A* **373**, 1–20 (2015).
31. Langelier, B., Wang, X. & Grandfield, K. Atomic scale chemical tomography of human bone. *Sci. Rep*. **7**, 39958 (2017).
32. Stoffers, A. et. al. Correlating atom probe tomography with atomic-resolved scanning transmission electron microscopy: example of segregation at silicon grain boundaries. *Microsc. Microanal*. **23**, 291–299 (2017).
33. Daculsi, G. & Kerebel, B. High-resolution electron-microscope study of human enamel crystallites — size, shape, and growth. *J. Ultrastruct. Res*. **65**, 163–172 (1978).
34. Beniash, E., Metzler, R. A., Lam, R. S. & Gilbert, P. U. Transient amorphous calcium phosphate in forming enamel. *J. Struct. Biol*. **166**, 133–143 (2009).
35. Robinson, C., Fuchs, P. & Weatherell, J. A. The appearance of developing rat incisor enamel using a freeze fracturing technique. *J. Cryst. Growth* **53**, 160–165 (1981).
36. Kirkham, J. et al. Self-assembling peptide scaffolds promote enamel remineralization. *J. Dent. Res*. **86**, 426–430 (2007).
37. Luder, H. U., Gerth-Kahlert, C., Ostertag-Benzinger, S. & Schorderet, D. F. Dental phenotype in Jalili syndrome due to a c.1312 dupC homozygous mutation in the *CNNM4* gene. *PLoS ONE* **8**, e78529 (2013).
38. Yamazaki, D. et al. Basolateral Mg^{2+} extrusion via CNNM4 mediates transcellular Mg^{2+} transport across epithelia: a mouse model. *PLoS Genet*. **9**, e1003983 (2013).
39. Nakano, Y. et al. A critical role of TRPM7 as an ion channel protein in mediating the mineralization of the craniofacial hard tissues. *Front. Physiol*. **7**, 258 (2016).
40. Aoba, T., Shimoda, S. & Moreno, E. C. Labile or surface pools of magnesium, sodium, and potassium in developing porcine enamel mineral. *J. Dent. Res*. **71**, 1826–1831 (1992).
41. Hubbard, M. J., Mangum, J. E., Perez, V. A., Nervo, G. J. & Hall, R. K. Molar hypomineralisation: a call to arms for enamel researchers. *Front. Physiol*. **8**, 546 (2017).

Methods

Consumables

Unless otherwise specified, all solutions were prepared using ultra-pure water (18.2 MΩ cm) dispensed from a Barnstead Nanopure UF+UV unit (Thermo-Fisher Scientific). Lactic acid ($C_3H_6O_3$; Mallinckrodt Chemicals); propionamide (98%), NaF (Sigma-Aldrich); NaH_2PO_4, Na_2HPO_4, HNO_3 (65 wt%), $Mg(NO_3)_2\cdot6H_2O$ (99%), $Ca(NO_3)_2\cdot4H_2O$ (99%, lot no. 86432), ethanol (VWR, Radnor, PA); $Ca(NO_3)_2\cdot4H_2O$ (99.98%, lot no. 61600281), formaldehyde (CH_2O) (Alfa Aesar); $(NH_4)_2HPO_4$ (99%, lot A0059707, Merck KGaA); PELCO liquid silver paint, graphite tape (Ted Pella); EPO-TEK 301 (Epoxy Technology); CarbiMet SiC grinding paper, Metadi supreme polycrystalline aqueous diamond polishing suspension, Microcloth polishing cloth (Buehler); and MM22 microtip coupons for FIB liftout (CAMECA Instruments).

Preparation of enamel sections

De-identified human premolars extracted for orthodontic reasons were kept in 10% buffered formalin at room temperature for 10 days, and at 4 °C thereafter. Before use, samples were rinsed with water and dried under a gentle stream of nitrogen gas ('rinsed and dried'). Samples were embedded in Epo-Tek 301 epoxy, sectioned along the buccal-lingual line, ground with SiC paper (600, 800, 1,200 grit), and polished on a Buehler Trident polishing cloth with polycrystalline diamond suspensions (3 µm, 1 µm), rinsed and dried. Some sections were treated with fluoride by immersion in 50 ml of aqueous NaF (250 mM, pH 8.4) under gentle agitation using an orbital shaker, at 37 °C for 24 h, then rinsed and dried. Some sections were exposed to lactic acid (250 mM, pH 4, for ~20 s), and subsequently rinsed and dried. Some sections, oriented such that rods (and crystallites) emerge approximately perpendicular to the surface, were positioned at an angle of roughly 45° with respect to the underlying lab bench surface. The sample was etched using a steady stream of droplets of lactic acid (250 mM, pH 4) flowed across the surface for a total of 30 s. Thereafter, the tooth sections were immediately immersed in ethanol, rinsed and dried.

Unless otherwise noted, samples were affixed to an aluminium stub using carbon tape, coated with AuPd (~25 nm) using a Denton Desk IV sputter coating system (Denton Vacuum). The surface of the sample was then grounded to the stub using colloidal silver paint.

Scanning electron microscopy (SEM)

SEM was performed using a Hitachi S4800-II or a Hitachi SU8030 (Hitachi High-Tech), both equipped with a cold cathode field emission electron gun, operated at an accelerating voltage of 5 kV and an emission current of 8,600 nA. Images were acquired using secondary electron contrast.

Transmission electron microscopy (TEM)

Lamellae were prepared from ground and polished, transverse sections of outer, buccal human enamel. A dual-beam FIB/SEM (FEI Helios NanoLab or FEI Strata 400) with a gallium liquid metal ion source (LMIS) operating at an accelerating voltage of 2–30 kV was used to prepare FIB samples for TEM. A ~200-nm-thick layer of protective carbon was deposited on a 2 µm × 15 µm area of interest, either by using the electron beam (5 kV, 1.4 nA) through decomposition of a phenanthrene precursor gas (FEI Helios Nanolab), or by selecting a similar area of interest using a permanent marker deposition method[42] (FEI Strata 400). On top of the carbon, a ~1-µm-thick protective platinum layer was deposited using the ion beam (30 kV, 93 pA) through decomposition of a (methylcyclopentadienyl)-trimethyl platinum precursor gas. Subsequently, two trenches were cut to allow for a roughly 2-µm-thick lamella of enamel. Next, the micromanipulator was welded onto the lamella, and the sample was cut loose from the bulk material. An in situ liftout of the sample was performed, and the lamella was welded onto a TEM half-grid. After thinning to about 40 nm in a sub-region of the lamella (5 kV, 81 pA), the section was cleaned at low voltage and current (2 kV, 28 pA) until a final thickness of 20–30 nm was achieved.

Scanning TEM (STEM) images were acquired on a JEOL JEM-2100F (JEOL USA, Peabody, MA), a JEOL GrandARM 300F, or an aberration-corrected FEI Titan Themis (FEI Co.) equipped with a monochromator, a side-entry double tilt liquid-nitrogen-cooled sample holder (Gatan 636; Gatan, Pleasanton, CA), and a cryogenically cooled anti-contamination device (for typical conditions see Supplementary Table 1). Image post-processing was performed as described in the Supplementary Information.

STEM-EDS

EDS spectrum images of regions of interest chosen in STEM-HAADF images were acquired on a windowless 100 mm^2 XMax^N 100TLE Silicon Drift Detector (SDD) with a solid angle of ~0.98 sr (Oxford Instruments NanoAnalysis) with a dwell time of 5 µs per pixel.

Cryo-STEM-EELS

EELS 2D spectrum images were acquired on a Titan Themis (FEI) using a K2 Summit direct electron detector in counting mode (Gatan). This direct electron detector with high quantum efficiency (DQE up to 80%) allowed simultaneous acquisition of all relevant inner shell ionization (core loss) edges at high energy resolution and low background levels despite the low dose required to minimize beam damage[14]. The entrance aperture was 5 mm, the energy dispersion 0.5 eV per channel. The beam current was 4.0–8.5 pA, the dwell time 2.5 µs per pixel. STEM-ADF images were recorded in parallel. Concentration maps were extracted by fitting and subtracting the pre-edge background with a linear combination of power laws and integrating the intensity under the EELS edge of interest. MCR analysis of the Mg $L_{2,3}$-edge region (Supplementary Fig. 6) was performed as described previously[43].

X-ray absorption spectroscopy (XAS)

XAS measurements were performed at the Spherical Grating Monochromator (SGM, 11ID-1) at the Canadian Light Source (Saskatoon, Saskatchewan, Canada), following a literature protocol[7]. Briefly, enamel from de-identified human third molars was ground into a powder using an agate mortar pestle and spread on graphite tape. Spectra of the Mg K-edge (1,303 eV) were recorded from −60 eV to −12 eV in steps of 2 eV, from −12 eV to −8 eV in steps of 0.5 eV, from −8 eV to 30 eV in steps of 0.1 eV, from 30 eV to 190 eV in steps of 0.2 eV, from 190 eV to 300 eV in steps of 0.3 eV, and from 300 eV to 400 eV in steps of 0.5 eV, with a constant dwell time of 2 s per step. Monochromator energy calibration was performed by setting the first absorbance maxima of the MgO reference sample spectra to 1,309.5 eV. X-ray fluorescence intensity was measured simultaneously with four solid-state silicon drift energy dispersive X-ray detectors (Amptek). Incident flux was measured by recording the current from a gold mesh upstream. The exit slit was adjusted and the undulator detuned to reduce flux to prevent saturation of X-ray fluorescence detectors when measuring concentrated reference samples. Between 1 and 7 scans were collected for each sample and averaged. No beam-induced changes were observed when comparing sequential spectra. The Mg X-ray fluorescence intensity was isolated from the total fluorescence intensity (which contained contributions from X-ray fluorescence from other elements and the scattered incident beam) using custom written code in Mathematica (Wolfram Research). For XANES spectra, see Supplementary Fig. 7.

Absorption data were normalized, background-subtracted using AUTOBK, and converted to k-space using Athena[44]. Edge energy (E_0) was set to the maximum of the first derivative of the absorption spectra. $\chi(k)$ data were weighted by k^2 and Fourier-transformed over a k-range of 2–9.5 Å^{-1}, applying a Hanning window with a sill width of 1 Å^{-1}. Theoretical photoelectron scattering amplitudes and phase shifts based on the crystal structures of dolomite[45], huntite[46], whitlockite[47] and hydroxyapatite[22] were calculated using FEFF6[48]. Shell-by-shell fitting of the EXAFS data was performed in R-space using Artemis[44]. An energy shift parameter (ΔE_0) was maintained constant for the scattering paths but allowed to vary between samples. The amplitude reduction factor ($S_0^2 = 0.8$) was determined on the basis of a fit to the dolomite, huntite and whitlockite spectra with coordination numbers constrained based on their respective crystal structures. Multiple scattering in the carbonate reference samples was accounted for following Reeder and co-workers[49]. Enamel and ACP EXAFS spectra were fitted using a model based on the Ca[II] site of hydroxylapatite, consisting of a single Mg-O and two Mg-P scattering paths[50,51]. To minimize the number of fitting parameters, the coordination number and σ^2 for the two Mg-P paths were constrained for each sample but allowed to vary between samples. For EXAFS spectra, see Supplementary Fig. 7. For fitted parameters, see Supplementary Table 3.

Atom probe tomography (APT)

Samples for APT were extracted ~10 µm below the external enamel surface on mid-coronal cervical sections of human premolars, using a Dual Beam SEM/FIB (Helios NanoLab; FEI), and following standard protocols[52,53]. Briefly, a 200-nm-thick layer of protective "FIB-Pt" was deposited using the electron beam (5 kV, 1.4 nA) on a 2 µm × 25 µm area of interest through decomposition of a (methylcyclopentadienyl) trimethyl platinum precursor gas. A thicker coating of FIB-Pt (~400 nm) was then deposited using the ion beam (30 kV, 93 pA). An angled cut was then made on either side of the Pt strap, and one end was cut free and attached to an in-situ manipulator (Omniprobe) using FIB-Pt. After cutting the final side free, 1–2 µm segments were attached to the top of silicon posts on the APT array with FIB-Pt. Tips were sharpened in the ion beam using annular mill patterns with progressively smaller inner and outer diameters (16–30 kV, 0.28–0.47 nA). The majority of contamination/gallium implantation was removed by a final cleaning step (2 kV, 0.25 nA).

APT analysis was performed using a LEAP 5000 XS (CAMECA Instruments) with a laser operating at a wavelength of 355 nm and a pulse frequency of 250 kHz, at a power of 40 pJ. The temperature in the analysis chamber was kept 25 K, the pressure $<10^{-8}$ Pa. The d.c. potential on the microtip was adjusted to maintain an evaporation rate of 0.005 ions per laser pulse. 3D reconstructions of the sample tips were made using the IVAS software package (CAMECA Instruments). Standard parameters were used for all reconstructions.

For representative APT spectra, see Supplementary Fig. 9. For peak identities and integration limits, see Supplementary Table 4, and for a comparison of the composition of human and rodent samples see Supplementary Table 6. In this manuscript, we analyse data from three reconstructions of fluoridated enamel, and two reconstructions of enamel that were not fluoridated by us. Inspection revealed that the cross-sections of 8 crystallites (8 fluoridated, 0 non-fluoridated) were fully contained in the reconstructions, and that of 14 crystallites (7/5) were partially contained. For 20 crystallites (15/5) we were able to extract 1D concentration profiles approximately normal to the midplane of the crystallite, using regions of interest (ROIs) that were defined manually in IVAS (CAMECA Instruments; Supplementary Fig. 12). For the remaining crystallites, too small a part was contained in the reconstruction, or it was not possible to deduce the orientation. Consequently, 1D profiles could not be extracted. 1D profiles were corrected for a homogeneous background. Additional reconstructions are given in Supplementary Figs. 10 and 11. As a visual aid, we animated our reconstruction of a 30-nm-thick slice through one enamel crystallite, showing individual ion positions in 3D space (Supplementary Video 1) and iso-concentration surfaces (Supplementary Video 2). We used the Matlab platform (The Mathworks) to render these videos.

Hydrothermal syntheses

Hydroxylapatite (OHAp, 0 at.% Mg) was synthesized following a literature protocol[54]. Briefly, 5 ml of an aqueous solution of 99.98% $Ca(NO_3)_2{\cdot}4H_2O$ (0.1 M, 0.5 mmol) was mixed with 5 ml aqueous solution of $(NH_4)_2HPO_4$ (0.06 M, 0.3 mmol). To the resulting suspension, 5 ml of an aqueous solution of propionamide (1 M, 5 mmol) was added. The pH was adjusted to 3 by addition of approximately 45 µl of aqueous HNO_3 (5 M), to give a clear, transparent solution. The solution was transferred to a PTFE-lined microwave digestion vessel and treated hydrothermally (heating ramp: 30 °C min^{-1}, final temperature of 180 °C held for 30 min), using a Milestone EthosEZ Microwave Digestion System (Milestone). The resulting precipitate was centrifuged and washed with deionized water (3 × 15 ml) and ethanol (3 × 15 ml), and dried in vacuo.

OHAp (0.22 at.% Mg) was synthesized as described above, but using 99% $Ca(NO_3)_2{\cdot}4\ H_2O$ instead of 99.98% $Ca(NO_3)_2{\cdot}4\ H_2O$.

OHAp (1.15 at.% Mg) was synthesized as follows: to 475 µl of an aqueous solution of 99.98% $Ca(NO_3)_2{\cdot}4\ H_2O$ (1 M, 0.475 mmol) was added 25 µl of an aqueous solution of $Mg(NO_3)_2{\cdot}6\ H_2O$ (1 M, 0.025 mmol). The solution was diluted to 5 ml overall volume with deionized water. A 5 ml aqueous solution of $(NH_4)_2HPO_4$ (0.06 M, 0.3 mmol) and a 5 ml aqueous solution of propionamide (1 M, 5 mmol) were added to give a suspension. The pH was adjusted to 3 by addition of approximately 45 µl aqueous HNO_3 (5 M) to give a clear, transparent solution. The solution was transferred to a PTFE-lined microwave digestion vessel and treated hydrothermally as described above. The resulting precipitate was centrifuged and washed with deionized water (3 × 15 ml) and ethanol (3 × 15 ml), and dried in vacuo to give OHAp (approximately 20 wt% by powder X-ray diffraction (PXRD)) and whitlockite (approximately 80 wt% by PXRD). Needle-shaped OHAp crystals could easily be differentiated from whitlockite platelets.

Phase identity and purity for all samples were confirmed by PXRD. The magnesium mole fraction was determined using ICP-MS.

Powder X-ray diffraction

PXRD patterns of OHAp were collected at 100 K on a STOE-STADI-P powder diffractometer (STOE Corporation) equipped with an asymmetric curved germanium monochromator (Cu Kα1 radiation, $\lambda = 1.54056$ Å) and a one-dimensional silicon strip detector (MYTHEN2 1K, DECTRIS). The line-focused Cu X-ray tube was operated at 40 kV and 40 mA. Powder was packed in a polyimide capillary (0.5 mm inner diameter) and intensity data were collected over an angular range of $2\theta = 10°–70°$, over a period of 10 min. The instrument was calibrated against an NIST silicon standard (640d). Data were processed and Rietveld refinement was performed using MDI Jade 2010 (Materials Data). Lattice parameters are reported in Supplementary Table 7.

Single crystal X-ray diffraction

Diffraction data for OHAp (0.22 at.% Mg) and OHAp (1.15 at.% Mg) were collected at a set temperature of 100 K using a Bruker Kappa APEX2 diffractometer (Bruker AXS) equipped with a Mo Kα ($\lambda = 0.71073$ Å) source. Single crystals (50 µm × 5 µm × 5 µm) were picked from powders and mounted with Paratone N on a cryo-loop. Diffraction patterns were indexed, refined, and integrated using SAINT of the APEX2 package (Bruker AXS). Using Olex2[55], the structure was solved with XT and refined with the ShelXL package using least squares minimization (Supplementary Fig. 14)[56]. Lattice parameters are reported in Supplementary Table 7 and Supplementary Fig. 15.

Inductively coupled plasma mass spectrometry (ICP-MS)

ICP-MS was carried out on a Thermo iCAP Q ICP-MS (Thermo Fisher Scientific). For phase-pure samples, powders were used as-received. For the sample containing whitlockite, the single crystal that was analysed by X-ray diffraction was used. Samples were dissolved in trace-metal-grade HNO_3 solution (0.1 M) in a metal-free tube. Trace-metal-grade HCl solution was used as a blank.

Density functional theory (DFT) calculations

DFT calculations were performed within the generalized gradient approximation (GGA) using the GGA-PBEsol (Perdew–Burke–Ernzerhof revised for solids) exchange-correlation functional[57] with the planewave pseudopotential code, Quantum ESPRESSO[58]. We used the ultrasoft pseudopotentials[59] taken from the PSLibrary[60]. A plane-wave cutoff of 60 Ry was used during the ionic and electronic relaxation steps. For the simulation of Mg-doped OHAp solid solutions, we employed a 2 × 2 × 1 supercell (with 352 atoms) in the monoclinic crystal structure ($P2_1/c$ symmetry). Our initial simulation performed on a pristine $Ca_{10}(PO_4)_6(OH)_2$ supercell serves as the reference. Additional simulations were performed, whereby 1, 2, 3 and 4 Ca atoms were substituted with Mg atoms. The atomic positions and the cell volume were relaxed until the Hellmann–Feynman forces were less than 2 meV Å^{-1} and components of the stress tensor were less than 0.1 kbar. The Brillouin zone integration was performed using a 1 × 1 × 1 Monkhorst–Pack *k*-point mesh (Γ-point calculation). Lattice parameters as calculated (at 0 K) and corrected to 298 K using the coefficient of thermal expansion reported by Babushkin and co-workers[61] are reported in Supplementary Fig. 15.

Lattice strain as a function of composition

Lattice parameters as a function of magnesium mole fraction were determined by XRD and DFT as outlined above (Supplementary Fig. 15). Lattice parameters as a function of carbonate weight fraction were obtained from Demier and co-workers[25] using WebPlotDigitizer[62]. Carbonate weight fractions were converted to mole fractions (X_C) using the stoichiometric models postulated by Deymier and co-workers[25]. The concentration-dependent lattice strain in the *a* and *c* directions was calculated from the lattice parameters and fitted with a linear model:

$$\varepsilon_i^j = n_i^j X_j + b_i^j$$

where ε_i^j denotes the strain in the *i* direction due to substitution with species *j*, X_j is the mole fraction of species *j*, n_i^j is the slope, and b_i^j is the intercept (Supplementary Figs. 16 and 17). Fit parameters are reported in Supplementary Table 8.

Finite-element modelling

Enamel crystallites were idealized as slabs with rectangular cross-section, oriented with the [001] direction parallel to the *z* axis and the [100] parallel to the *x* axis. Mg and carbonate concentrations were modelled as continuous, 2D distributions (Supplementary Fig. 18), chosen to represent experimental 1D concentrations profiles in the *x* direction (Supplementary Fig. 13). For simplicity, the contributions of fluoride and of sodium ions in excess of those needed to charge-balance carbonate ions were ignored, as were contributions from surface free energies. All modelling was performed using COMSOL Multiphysics.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request. Source data are provided with this paper.

Code availability

This manuscript primarily made use of commercial (IVAS, Origin, Matlab, MDI Jade, APEX2, Thermo Scientific Qtegra ISDS, COMSOL Multiphysics, TEM Imaging and Analysis, DigitalMicrograph, AZtec, Adobe Illustrator) and freely available (DEMETER, OLEX[2], SHELX, Quantum ESPRESSO, Cornell Spectrum Imager, ImageJ) software packages for acquisition, processing and visualization of data. MCR was performed using custom code using the Matlab mcr.m package from the Eigenvector Research PLS_toolbox, as described elsewhere[43]. In addition, custom code written for the Mathematica and Matlab environments was used for file conversions, plotting and visualization. This code is available from the corresponding author upon reasonable request.

42. Park, Y. C. et al. Use of permanent marker to deposit a protection layer against FIB damage in TEM specimen preparation. *J. Microsc.* **255**, 180–187 (2014).
43. Zachman, M. J., Tu, Z. Y., Choudhury, S., Archer, L. A. & Kourkoutis, L. F. Cryo-STEM mapping of solid-liquid interfaces and dendrites in lithium-metal batteries. *Nature* **560**, 345–349 (2018).
44. Ravel, B. & Newville, M. ATHENA, ARTEMIS, HEPHAESTUS: data analysis for X-ray absorption spectroscopy using IFEFFIT. *J. Synchrotron Radiat.* **12**, 537–541 (2005).
45. Antao, S. M., Mulder, W. H., Hassan, I., Crichton, W. A. & Parise, J. B. Cation disorder in dolomite, $CaMg(CO_3)_2$, and its influence on the aragonite + magnesite ↔ dolomite reaction boundary. *Am. Mineral.* **89**, 1142–1147 (2004).
46. Dollase, W. A. & Reeder, R. J. Crystal-structure refinement of huntite, $CaMg_3(CO_3)_4$, with X-ray powder data. *Am. Mineral.* **71**, 163–166 (1986).
47. Calvo, C. & Gopal, R. Crystal structure of whitlockite from Palermo Quarry. *Am. Mineral.* **60**, 120–133 (1975).
48. Rehr, J. J. & Albers, R. C. Theoretical approaches to X-ray absorption fine structure. *Rev. Mod. Phys.* **72**, 621–654 (2000).
49. Reeder, R. J., Lamble, G. M. & Northrup, P. A. XAFS study of the coordination and local relaxation around Co^{2+}, Zn^{2+}, Pb^{2+}, and Ba^{2+} trace elements. *Am. Mineral.* **84**, 1049–1060 (1999).
50. Holt, C. et al. Preparation of amorphous calcium-magnesium phosphates at pH 7 and characterization by x-ray absorption and fourier transform infrared spectroscopy. *J. Cryst. Growth* **92**, 239–252 (1988).
51. Harries, J. E., Hukins, D. W. L., Holt, C. & Hasnain, S. S. Conversion of amorphous calcium phosphate into hydroxyapatite investigated by EXAFS spectroscopy. *J. Cryst. Growth* **84**, 563–570 (1987).
52. Larson, D. J., Prosa, T. J., Ulfig, R. M., Geiser, B. P. & Kelly, T. F. *Local Electrode Atom Probe Tomography: A User's Guide* (Springer Science and Business Media, 2013).
53. Thompson, K. et al. In situ site-specific specimen preparation for atom probe tomography. *Ultramicroscopy* **107**, 131–139 (2007).
54. Qi, M.-l., Xiao, G.-y. & Lu, Y.-p. Rapid hydrothermal synthesis of submillimeter ultralong flexible hydroxyapatite fiber using different pH regulators. *Acta Metall. Sinica Engl. Lett.* **29**, 609–613 (2016).
55. Dolomanov, O. V., Bourhis, L. J., Gildea, R. J., Howard, J. A. K. & Puschmann, H. OLEX2: a complete structure solution, refinement and analysis program. *J. Appl. Crystallogr.* **42**, 339–341 (2009).
56. Sheldrick, G. M. Crystal structure refinement with SHELXL. *Acta Crystallogr. C* **71**, 3–8 (2015).
57. Perdew, J. P. et al. Restoring the density-gradient expansion for exchange in solids and surfaces. *Phys. Rev. Lett.* **100**, 136406 (2008).
58. Giannozzi, P. et al. QUANTUM EXPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys. Condens. Matter* **21**, 395502 (2009).
59. Vanderbilt, D. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Phys. Rev. B* **41**, 7892–7895 (1990).
60. Dal Corso, A. Pseudopotentials periodic table: from H to Pu. *Comput. Mater. Sci* **95**, 337–350 (2014).
61. Babushkin, O., Lindbäck, T., Holmgren, A., Li, J. & Hermansson, L. Thermal expansion of hot isostatically pressed hydroxyapatite. *J. Mater. Chem.* **4**, 413–415 (1994).
62. Rohatgi, A. WebPlotDigitizer Version 4.2 https://automeris.io/WebPlotDigitizer (2019).

Acknowledgements This work was supported in part by the National Institute of Health–National Institute of Dental and Craniofacial Research (NIH-NIDCR R03 DE025303-01, R01 DE025702-01), the National Science Foundation (DMR-1508399), the NSF Platform for the Accelerated Realization, Analysis and Discovery of Interface Materials (PARADIM) under Cooperative Agreement no. DMR-1539918, and the University of Virginia. K.A.D. was in part supported by a 3M fellowship. The Canadian National Sciences and Engineering Research Council in part supported L.M.G. K.A.D. and M.J.C. were supported in part by the Northwestern University Graduate School Cluster in Biotechnology, Systems and Synthetic Biology, which is affiliated with the Biotechnology Training Program. L.S. was supported by a Deutsche Forschungsgemeinschaft research fellowship (STE2689/1-1). This work made use of the following core facilities operated by Northwestern University: NUCAPT, which received support from NSF (DMR-0420532), ONR (N00014-0400798, N00014-0610539, N00014-0910781 and N00014-1712870), and the Initiative for Sustainability and Energy at Northwestern University (ISEN); MatCI; NUANCE and EPIC, which received support from the International Institute for Nanotechnology (IIN), the Keck Foundation, and the State of Illinois, through the IIN; IMSERC; the Jerome B. Cohen X-Ray Diffraction Facility; QBIC, which received support from NASA Ames Research Center (NNA06CB93G). NUCAPT, MatCI, NUANCE and EPIC were further supported by the MRSEC programme (NSF DMR-1720139) at the Materials Research Center; NUCAPT, NUANCE, EPIC and IMSERC were also supported by the Soft and Hybrid Nanotechnology Experimental (SHyNE) Resource (NSF ECCS-1542205). This work made use of the Cornell Center for Materials Research (CCMR) Shared Facilities supported through the NSF MRSEC Program (no. DMR-1719875). The Titan Themis 300 was acquired through NSF-MRI-1429155, with additional support from Cornell University, the Weill Institute, and the Kavli Institute at Cornell. This work made use of the Rivanna cluster maintained by the Advanced Research Computing Services at the University of Virginia. Portions of this work were performed at the Canadian Light Source (CLS), which received support from The Natural Sciences and Engineering Research Council of Canada, the National Research Council of Canada, the Canadian Institutes of Health Research, the Province of Saskatchewan, Western Economic Diversification Canada, and the University of Saskatchewan. The authors thank P. Akers, M. Stohle and G. Borden for providing de-identified human premolars, and C. Malliakas, K. MacRenaris, M. Thomas and especially K. Rice for technical support.

Author contributions The experiments and simulations were designed by K.A.D., P.J.M.S., L.S., P.V.B., J.M.R., D.J. STEM experiments were performed by P.J.M.S., B.H.G. and M.J.Z., with additional help in analysis and simulations provided by M.J.Z. and L.F.K. APT data were collected and analysed by K.A.D., with D.J. assisting the analysis. The finite-element model was developed by K.A.D. and D.J. using DFT calculations performed by P.V.B., and synthetic Mg-rich OHAp crystals made and analysed by L.S. XAS data were collected and analysed by L.M.G. and M.J.C. K.A.D., P.J.M.S., B.H.G., P.V.B., M.J.Z., L.F.K., J.M.R. and D.J. were all involved in preparing the manuscript.

Competing interests The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2433-3.

Correspondence and requests for materials should be addressed to D.J.

Reprints and permissions information is available at http://www.nature.com/reprints.

nature research

Corresponding author(s): Derk Joester

Last updated by author(s): 3/30/2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	XAS: data was collected at the using software provided by the beamline. APT: data was collected using IVAS (CAMECA Instruments, Inc., Madison, WI). XRD: powder diffraction data was collected using WinXPOW (STOE Corporation, Chicago, IL); single-crystal diffraction patterns were collected using APEX2 (Bruker AXS, Inc., Madison, WI). ICP-MS data was collected using Thermo Scientific™ Qtegra™ Intelligent Scientific Data Solution™ (Thermo Fisher Scientific, Inc.; Waltham, MA). STEM images were acquired using TEM Imaging and Analysis (FEI Company; Hillsboro, OR) or DigitalMicrograph (Gatan Inc., Pleasanton, CA). EELS data was acquired using DigitalMicrograph. EDS data was acquired using AZtec (Oxford Instruments NanoAnalysis, Concord, MA).
Data analysis	XAS: Data was processed using ATHENA, ARTEMIS, FEFF6, and AUTOBK that are all part of the DEMETER system (J. Synchrotron Rad. 2005, 12, 537–541; https://bruceravel.github.io/demeter/). Data files were converted, and spectra and fits plotted using custom code written for the Mathematica (Wolfram Research, Inc., Champaign, IL) and Matlab (The MathWorks Inc., Natick, MA) environments. APT: Data was analyzed and visualized using IVAS (CAMECA Instruments, Inc., Madison, WI). Spectra were plotted and videos rendered using Matlab. XRD: Rietveld refinement of PXRD data was performed using MDI Jade 2010 (Materials Data, Inc., Livermore, Ca). Singe crystal XRD structures were solved using the APEX2 package (Bruker AXS, Inc., Madison, WI), Olex2 (OlexSys Ltd, Durham, UK), and the SHELX package (Acta Cryst. 2008, A64, 112-122; http://shelx.uni-goettingen.de). Processed data was plotted using Matlab (The MathWorks Inc., Natick, MA) and Origin 2018 (Originlab Corp. Northhampton, MA). Data was extracted from the literature using WebPlotDigitizer (https://automeris.io/WebPlotDigitizer) and plotted using Matlab. Finite Element Modeling (FEM) was performed and visualized using COMSOL Multiphysics® (COMSOL, Inc., Burlington, MA). Some data was plotted using Matlab. DFT calculations made use of the Quantum ESPRESSO suite of open source codes (J.Phys. Condens. Matter 2009, 21, 395502; J. Phys. Condens. Matter 2017, 29, 465901; https://www.quantum-espresso.org). Data was plotted using Matlab. ICP-MS data was analyzed using Thermo Scientific™ Qtegra™ Intelligent Scientific Data Solution™ (Thermo Fisher Scientific, Inc.;

Waltham, MA).
STEM image registration performed using custom code in Python as described in Ultramicroscopy 2018, 191, 56-65. EELS data was processed and analyzed using the Cornell Spectrum Imager plugin (Microsc. Microanal. 2012, 18, 667-675; http://spectrumimager.com) for ImageJ (BMC Bioinformatics 2017, 18, 529; https://imagej.net/Welcome). MCR was performed in using the Matlab mcr.m package from the Eigenvector Research PLS_toolbox, as described in Nature 2018, 560, 345–349. EDS data was processed using AZtec (Oxford Instruments NanoAnalysis, Concord, MA).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Raw data that support the findings of this study are available from the corresponding author upon reasonable request. Source data for Figures 2 [b,c,d,e,f,i], 3 [e,f], 4c, S4 [b,c], S5f, S6a, S7 [a,b,c], S13, S14, S15, S16, and S17 are provided as part of the Supplemental Information.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size was not predetermined because this was an exploratory (pilot) study and we did not know what to expect. Sampling continued until three samples for the NaF treatment group and two samples from the untreated group were found to be of sufficient quality for further analysis (see exclusion criteria).
Data exclusions	A total of 23 samples ('tips') were prepared for analysis by APT. 18 of these represented native samples, and 5 represented samples treated with NaF. During initial quality control, data sets with unsatisfactory voltage history were excluded. Small data sets (<5-7 M hits) in which the likelihood of finding crystallites with cross sections that are fully included is low were also excluded.
Replication	Herein, we analyze three APT data sets collected from enamel after treatment with aqueous NaF (yield 60%, Table S5, Fig. S10a-c), and two data sets that were not treated (yield 11%, Table S5, Fig. S10d,e). Within theses 5 tips, we identified 20 crystallites that all share the same core-shell architecture in which the core itself has sandwich structure. We further observed the core-shell structure in hundreds of crystallites by STEM, confirming the observation of others using electron optical imaging. Finally, we observed etching of the core by SEM imaging for a large number of crystallites. We therefore propose that the APT findings are far more general than the relatively small sample suggests.
Randomization	There was no randomization because the objective was not to compare between treatment groups, but to establish fundamental properties of baseline samples.
Blinding	Blinding was not required, as this study did not try to establish differences between treatment groups.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Article

Abrupt increase in harvested forest area over Europe after 2015

https://doi.org/10.1038/s41586-020-2438-y

Received: 17 May 2019

Accepted: 23 April 2020

Guido Ceccherini[1]✉, Gregory Duveiller[1], Giacomo Grassi[1], Guido Lemoine[2], Valerio Avitabile[1], Roberto Pilli[1] & Alessandro Cescatti[1]

Forests provide a series of ecosystem services that are crucial to our society. In the European Union (EU), forests account for approximately 38% of the total land surface[1]. These forests are important carbon sinks, and their conservation efforts are vital for the EU's vision of achieving climate neutrality by 2050[2]. However, the increasing demand for forest services and products, driven by the bioeconomy, poses challenges for sustainable forest management. Here we use fine-scale satellite data to observe an increase in the harvested forest area (49 per cent) and an increase in biomass loss (69 per cent) over Europe for the period of 2016–2018 relative to 2011–2015, with large losses occurring on the Iberian Peninsula and in the Nordic and Baltic countries. Satellite imagery further reveals that the average patch size of harvested area increased by 34 per cent across Europe, with potential effects on biodiversity, soil erosion and water regulation. The increase in the rate of forest harvest is the result of the recent expansion of wood markets, as suggested by econometric indicators on forestry, wood-based bioenergy and international trade. If such a high rate of forest harvest continues, the post-2020 EU vision of forest-based climate mitigation may be hampered, and the additional carbon losses from forests would require extra emission reductions in other sectors in order to reach climate neutrality by 2050[3].

Forests provide a series of both tangible and intangible services to society and to human well-being, ranging from the production of raw materials and regulation of water flows to the protection of soils and conservation of biodiversity[4]. In the countries that form the EU, forests account for approximately 38% of the total land surface, out of which more than 95% are managed[1] with practices that vary broadly across countries[5,6]. Emerging wood markets driven by the bioeconomy—economic activities that use renewable biological resources to produce food, materials and energy—are challenging the current balance between wood demand and the need to preserve key ecosystem services[7]. In particular, in recent decades forests are increasingly considered to be a key asset for meeting climate mitigation targets[2]. Despite the mixed biophysical impacts of forests on climate[8–10], carbon sequestration by forests remains the most important negative climate forcing provided by forests at the global level[11]. In addition, further climate mitigation by forests may come from the increasing use of wood and wood-based residues for material and energy substitution, respectively[12].

On the policy side, the conservation and expansion of the forest carbon sink is an important element in the Paris Agreement[13], as these activities are expected to help countries to reach their individual mitigation goals and globally to achieve the required balance between anthropogenic greenhouse gas emissions and removals in the second half of the century[3]. Similarly, according to the recent European Green Deal[14], the EU's forested area needs to improve, both in quality (biodiversity and management) and in area, to reach climate neutrality and a healthy environment.

The amount of carbon sequestered by forest carbon sinks in the EU has remained stable over the last 25 years and currently offsets about 10% of total EU greenhouse gas emissions[15]. Most of this sink occurs in the living biomass, directly reflecting the difference between forest growth and forest harvest, mortality and natural disturbances. The rate of forest harvest is, therefore, a key parameter in forest management as it largely controls the forest carbon budget[16,17] and also affects ecosystem services such as the conservation of biodiversity, soils and water resources. In recent decades, harvested volumes in Europe's forests have been substantially lower than net annual growth[18], resulting in an increasing carbon stock. Given the fundamental relevance of the harvest rate, timely, consistent and robust assessments of the spatial patterns and temporal trends of the harvest rate are required in order to inform management policies and track economic and environmental progress towards a sustainable bioeconomy. However, official annual forest-harvest statistics typically do not cover the most recent years, their estimates are usually provided at a somewhat coarse spatial scale (by national or regional administrative units) and in some cases they are not regularly updated or are incomplete[19,20].

Currently, the combination of high-resolution satellite records and cloud-computing infrastructures that can handle 'big data' provides a complementary asset for quantifying harvested forest area that is independent from official statistics and overcomes some of the limitations of national inventories. Using such datastreams and information technologies, we assessed the recent changes (2004–2018) in harvested forest area based on the Hansen maps of Global Forest Change

---

[1]Bio-Economy Unit, European Commission Joint Research Centre, Ispra, Italy. [2]Food Security Unit, European Commission Joint Research Centre, Ispra, Italy.
✉e-mail: guido.ceccherini@ec.europa.eu

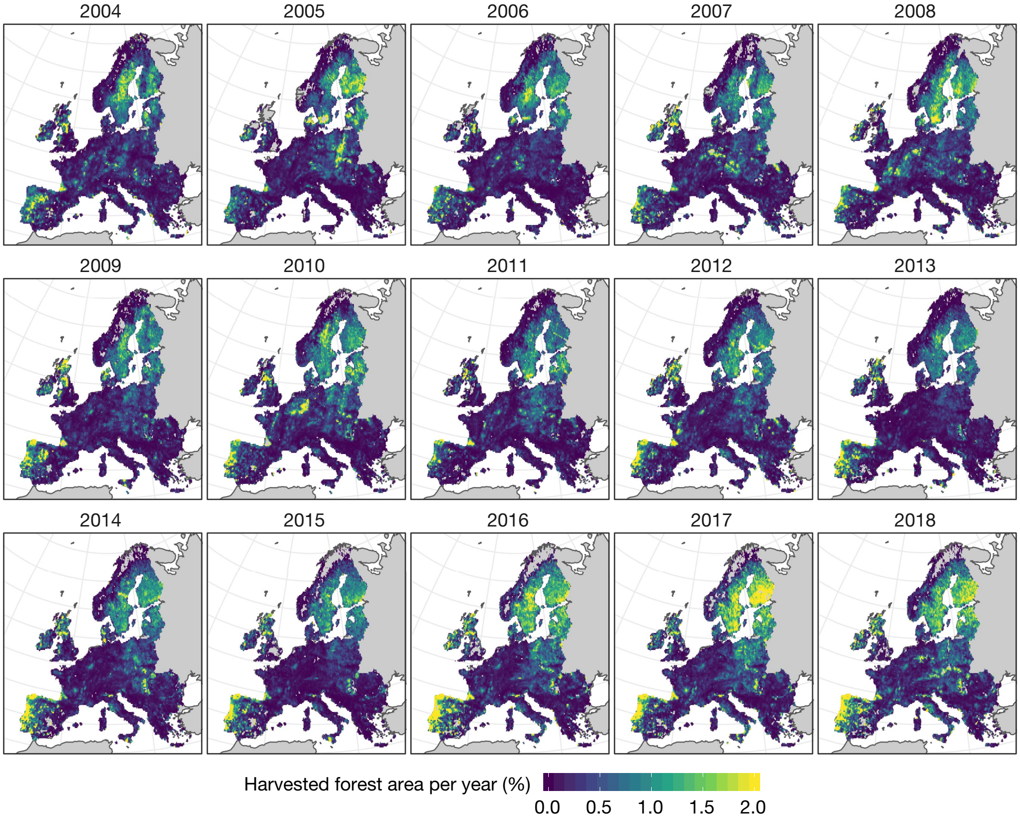


Fig. 1 | Harvested forest area per year. Percentage of harvested forest area (expressed as the relative amount of forest area affected by management practices) per year in a 0.2° grid cell, excluding forest losses due to fires and major windstorms and areas with sparse forest cover. For the generation of this map, land areas were classified only as forests when the tree cover exceeded a 20% threshold, uniformly throughout EU26, whereas the rest of the analysis was performed on the basis of a country-based tree-cover threshold as explained in Methods. Grey areas represent countries not included in the analysis. Map generated using GEE[22].

(GFC)[21], a map product with a 30-m resolution based on Landsat satellite data, which provides yearly estimates of tree cover and tree-cover loss (details in Methods section 'Forest mapping'). This evidence-driven assessment targets three questions: (1) whether, following the recent boost in the bioeconomy, the area of harvested forests is changing throughout the EU, and if so in which countries and to what degree; (2) which forests, in terms of biomass and plant cover type, show the largest changes in harvested rate; and (3) whether the modality of forest management in the EU is changing in terms of the size of harvested forest patches.

Here we estimate the changes in forest cover across 26 EU countries—including the UK and excluding Cyprus and Malta (herein referred to as EU26)—using the GFC maps implemented in Google Earth Engine[22], a big data Earth observation platform that enables seamless parallel computing and geospatial operations (details in Methods section 'Cloud-computing platform: Google Earth Engine'). Losses owing to forest fires and major windstorms (details in Methods section 'Spatial aggregation and major windstorm removal') are factored out. We assume that the annual loss in forest cover detected by the GFC maps is a reasonable proxy for the harvested forest area, because we remove losses related to fires or major windstorms. We note that the GFC dataset is sensitive to clear-cuts instead of the actual wood harvest, which can be complemented by thinning operations that may not be seen by the satellite—such as when the change in crown cover is not large enough to be detected.

Validation using a sample of high-resolution data (details in Methods section 'Validation of the GFC maps with high-resolution imagery') confirms the capacity of the GFC maps to detect forest loss, even though uncertainties are lower in some years compared to others, (for example, 2017 has lower uncertainty than 2012) and also lower for large patches (forest patch size greater than 0.27 ha) than in fragmented areas (patch size less than 0.27 ha) (Supplementary Fig. 1). The classification accuracy is particularly high (more than 82% correct detection) for patches larger than 4.5 ha, representing more than 60% of the detected harvested area in EU26. Henceforth, we refer to the forest-loss area as the harvested area.

In answering the first question, our results show that the intensity in harvest, defined here as the percentage of harvested forest area per year, was very stable in magnitude and spatial pattern across most EU26 countries from 2004 to 2015 (Fig. 1). Conversely, we observed a sudden increase in the mean value for the years 2016–2018: 43% with respect to the mean of the years 2004–2015 and 49% with respect to the mean of the years 2011–2015, with particular contributions from large EU domains such as the regions of Finland, Sweden, Lithuania, Latvia, Estonia and Poland, and the western part of the Iberian Peninsula. We acknowledge the uncertainty and the potential bias of the GFC maps, and in particular variations in the availability of observational data before and after 2012, owing to the frequency of Landsat acquisitions (see Methods section 'Forest mapping'). Nonetheless, we consider our findings reliable because abrupt changes in harvested forest area occurred in 2016–2018. We argue that these recent variations in harvested forest areas are due to changes in management and not to increased rates of natural disturbances from windstorms or fires, as these natural disturbances have been factored out from the analysis. This striking rise in harvested forest area is particularly marked in countries that have relevant forestry-related economic activities (for example, the bioenergy sector, paper industries), such as Sweden, Finland, Poland, France, Latvia, Portugal and Estonia. Although an

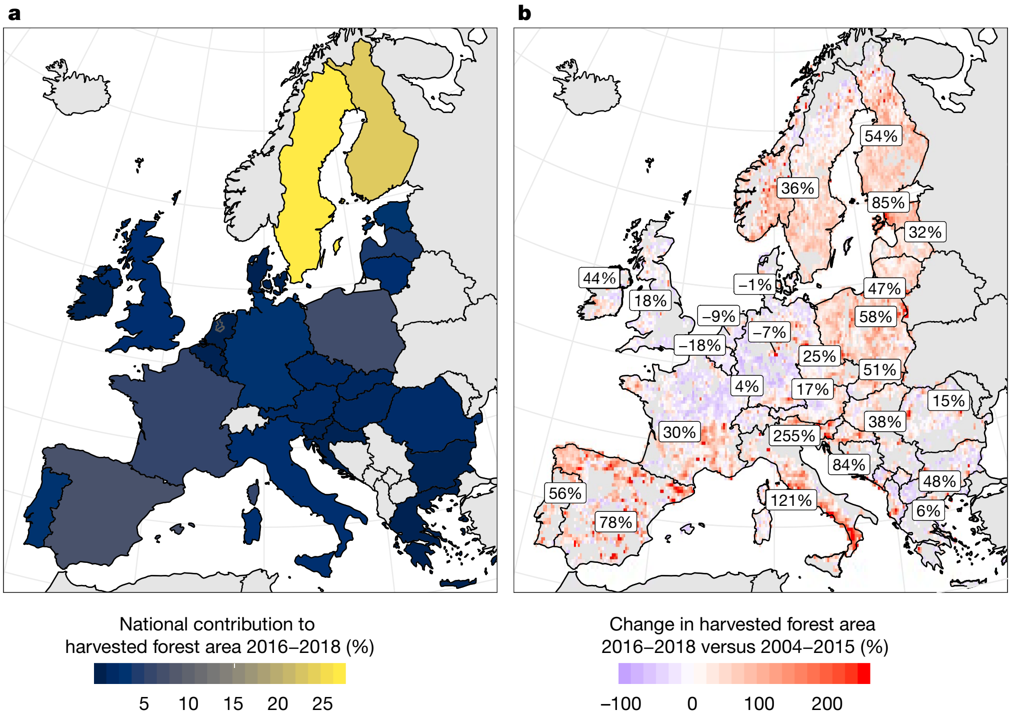


Fig. 2 | Spatial statistics of European harvested forest area. a, Percentage national contribution to the total harvested forest area of EU26 during 2016–2018. **b**, Percentage variation of European harvested forest area within each 0.2° × 0.2° grid cell, for 2016–2018 versus 2004–2015 (labels refer to aggregated national values). Grey areas represent countries not included in the analysis. Maps generated using GEE[22].

increased fraction of mature forests in the EU[18] is expected to drive a moderate increase in harvest rate in the coming decades[23], the magnitude and speed of change observed in 2016–2018 instead suggests an increase in wood demand and/or a change in forest management[24].

The largest share of variation in harvested forest area during 2016–2018 compared to 2004–2015 among the 26 EU countries was recorded in Sweden and Finland, which together accounted for more than 50% of the total increase in harvested area observed in recent years (Fig. 2a). Poland, Spain, France, Latvia, Portugal and Estonia accounted for about 30% in total. Needleleaf forests accounted for more than 50% of the detected harvested area in the 26 EU countries according to the European Space Agency (ESA) GlobCover global map on forest type[25], in agreement with the Eurostat report[26] (Extended Data Fig. 4). The analysis of the percentage variation (Fig. 2b) of the annual harvested forest area during 2016–2018 compared with the reference period (2004–2015) shows a general increase, with exceptions in Belgium, the Netherlands, Denmark and Germany, which show minor negative variations. The variation in harvested areas within each 0.2° × 0.2° grid cell confirms a widespread increase in harvested areas in Finland, Sweden, Latvia, Lithuania, Estonia, Poland, and the Iberian peninsula.

The assessment of the rate of forest harvest was quantified in terms of biomass loss by combining the GFC layer with a global map of above-ground biomass (AGB) in living trees for the year 2010, estimated from Earth observation data[27] (details in Methods section ‘Above-ground biomass analysis’). Results show that the patterns in biomass loss (Extended Data Fig. 8 and Supplementary Fig. 4) strongly resemble those of harvested area (Figs. 1, 2a). The increase of annual harvested forest biomass for the period 2016–2018 with respect to 2011–2015 is 69%, higher than the increase in harvested area during the same period. This implies that the areas harvested in the most recent years were characterized by a higher biomass density than those harvested in the reference period.

The 43% increase in annual harvested forest area observed for the years 2016–2018 relative to 2004–2015 was also accompanied by an increase in forest losses owing to natural disturbances from fires and windstorms, although these events were not included in the harvest-area statistics we report. An exceptional number of fires (an approximately 210% increase) were detected for the years 2016–2018 compared with the average number of fires observed during the 2004–2015 period (Fig. 3a). Major windstorms exhibited a rise of the order of 90%, especially in 2018, although the areas hit in 2016–2017 were generally smaller than those hit in 2005, 2007 and 2010.

The analysis of the time series of harvested forest area was carried out at EU26 country level and compared with existing statistics on harvested volume from FAOSTAT, further corrected to account for possible inconsistencies[17]. For this analysis we normalized the harvested volume to enable a comparison with harvested forest area (Extended Data Fig. 6). Overall, on the basis of a country-level analysis, we can conclude that remote-sensing estimates of harvested area are consistent with the statistics for harvested volume. Where inconsistencies were detected, country-specific circumstances—generally independent of the approach we propose here—were identified (details in Methods section ‘Harvested forest area at the country level and comparison with official harvest statistics’).

The second question we want to address is which forests—in terms of biomass and type—are undergoing the largest changes in their harvested area. Across EU26, we computed the average harvested forest area for five different biomass-density classes and the three major forest types (Fig. 3b). The analysis was carried out also for four selected countries (Supplementary Figs. 5 and 6): the two countries with the largest harvested areas (Sweden and Finland), one representative country in central Europe (Poland) and one country in southern Europe (Italy). Generally, the largest increase in harvested area during the period 2016–2018 occurred in needleleaf forests, followed by mixed and broadleaf forests, and the largest increase in the percentage of harvested area occurred in regions with 50–200 t ha^{-1} of biomass. The patterns of harvested biomass are different for different countries, reflecting the variability of forest types and management strategies

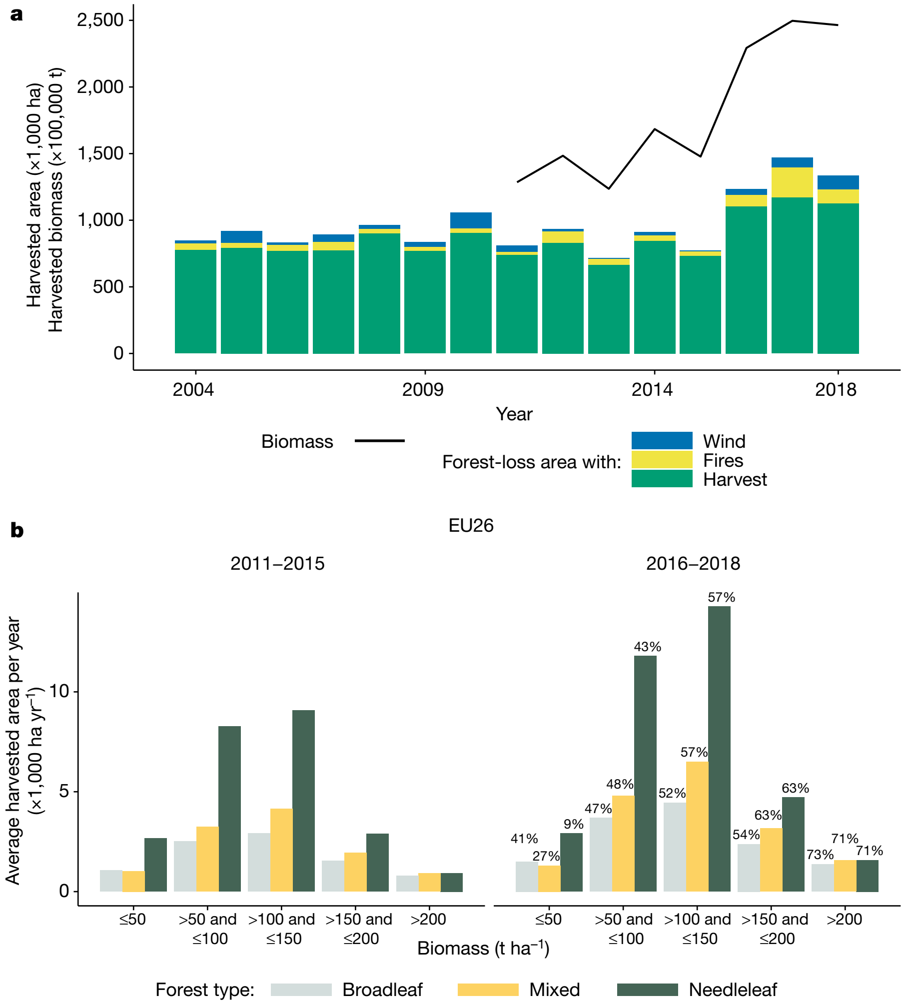


Fig. 3 | Temporal trends of forest harvests. a, Time series of forest biomass and area loss due to forest fires, major windstorms and harvested. **b**, Mean yearly harvested area for five biomass-density classes for the periods 2011–2015 (left) and 2016–2018 (right) for EU26, by forest type. Percentage labels at right show the variation (all are increases) for 2016–2018 compared with the reference period 2011–2015 for each biomass class.

across EU26. Both Finland and Sweden show a peak in harvested area for needleleaf forests with biomass density in the range 50–150 t ha^{-1}, whereas in Poland and Italy the maximum harvest values occur in mixed and broadleaf forests, respectively, that have higher biomass density (100–200 t ha^{-1}). This distribution of harvested area reflects the lower biomass stock of forest in the northern European countries compared with those in central Europe and also reflects the prevalence of broadleaves in southern Europe.

Taking advantage of the high spatial and temporal resolution of satellite records, we produced country statistics for the temporal trends of the size of harvested forest patches (that is, the median gap size), and the corresponding percentage variation in the median harvested patch size between 2004 and 2018 (Fig. 4). This analysis addresses our final question regarding ongoing changes in spatial patterns of harvested forest area. The size of harvested patches depends on the topography and silvicultural practices of the country, with larger patches observed in the case of massive clear-cuts, and smaller patches seen for group selection (in which groups or small patches of harvested area are created by the removal of adjacent trees) and shelterwood (in which young trees are grown under the shelter of older trees removed by successive cuttings) systems. The size of harvested patches may affect the impact of forest management on the provision of ecosystem services: generally, larger patches have stronger effects on ecosystems through habitat disruption, soil erosion and water regulation[28,29].

Satellite observations reveal that, overall, the median patch size has increased by 34% across EU26 (on the basis of the mean of the percentage changes of the individual EU26 countries for the years 2016–2018 compared with 2004–2015, weighted by total national forest area). The majority of this increase is attributable to large forest patches (>7.2 ha) (Extended Data Fig. 5). In 21 out of the 26 EU countries, the size of the harvested patches increased by more than 44% between the studied years. Portugal and Italy exhibit an abrupt rise in the median patch size for the period 2016–2018 compared with 2004–2015 (more than 100%). Also, the median patch size is substantially larger in Finland, Sweden, the UK and Ireland than in central or southern EU26 countries.

Exploring the reasons for the recent increase in harvested area, we identify three potential drivers: the ageing of European forests, an increase in salvage logging (owing to natural disturbances), and variations in socio-economic context, such as market demand and policy frameworks. Although harvest volumes can increase because of forest ageing[23], according to the most recent statistics[30] this cannot explain more than 10% of the observed increase in harvest area (details in Methods section 'Potential drivers of change in harvested forest area'). Moreover, the abrupt increase in harvested area as detected from satellite records is not consistent with the gradual trend expected from the effect of ageing. Additionally, although natural disturbances (such as forest fires, salvage logging after major windstorms and insect outbreaks) have affected inter-annual variations and trends, they have been factored out from the analysis. Thus, the socio-economic context

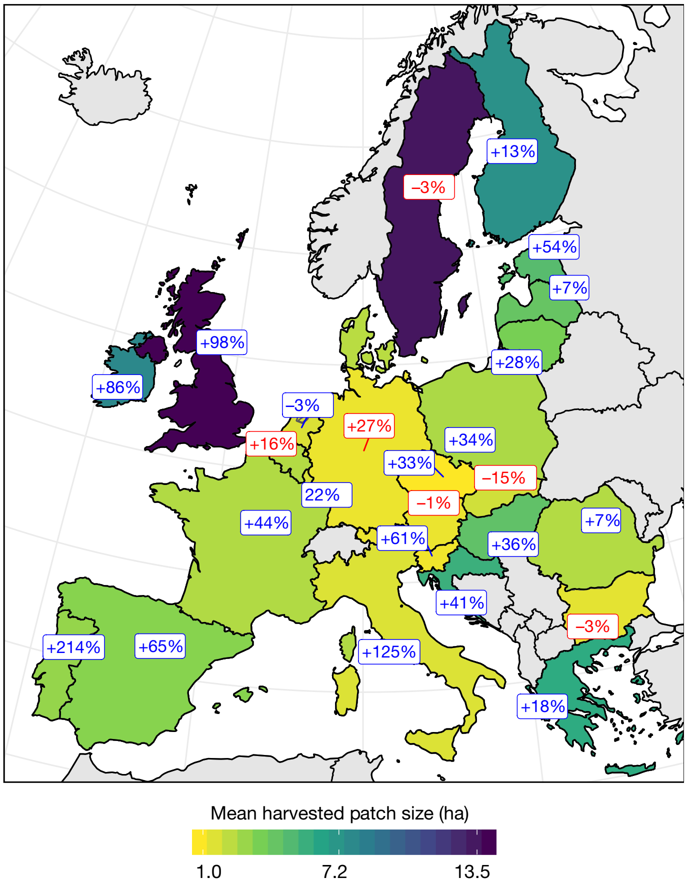


Fig. 4 | Mean harvested-patch size and recent change by country. Mean forest harvested patch size and the percentage variation for 2016–2018 compared with 2004–2015. The colour of the label indicates the agreement in sign between the variation in patch size and the total harvested forest area (red when in opposition, blue when in agreement with the harvest variations given in Fig. 2b). Grey areas represent countries not included in the analysis. Map generated using GEE[22].

and policy framework are most probably the most important drivers of harvest area increase, even if a causal connection is difficult to prove and quantify[31]. Although the effect on the harvest rate from a socio-economic stimulus or policy may vary from one country to another (including country-specific patterns of import and export), all economic indicators of wood demand and market (that is, FAOSTAT[30], Eurostat[32] and UNECE[33]) confirm a substantial expansion of the forest sector during the last years (details in Methods section 'Potential drivers of change in harvested forest area'). For example, the output of forestry and connected secondary activities (Extended Data Fig. 7) increased by 13% in 28 EU countries from 2012 to 2016 (as of the years of interest, thus including the UK). This is possibly linked to new legislation (at both EU and country levels) promoting the use of wood in the context of the bioeconomy[34], in particular in the use of renewable energy[35], which has been criticized for the potential impact on global forests[36].

Overall, our analysis shows that Earth observation can provide timely, independent, transparent and consistent monitoring of harvested forest areas across large geographical areas. Complementing national forest inventories with Earth observation has several benefits: (1) it increases transparency because governments or civil society (such as research centres and universities) can better track forest management, both spatially and temporally; (2) it supports the calculation of spatially explicit estimates of greenhouse gas emissions and removals, as required in recent EU land-related legislation[37]; (3) it enables increasing frequency of assessments, facilitating early warnings and timely policy responses; and (4) it assists in validating official statistics by enabling independent checks.

Our methodology, built on the large body of literature regarding the use of satellite remote sensing in the assessment of deforestation[38–40], was developed to deal with the specificity of forest management (such as different management types, no land usage change) and is thus a useful tool supporting the sustainable management of forests[41]. In the future, the interoperability of the NASA Landsat satellite with the ESA Copernicus Sentinels mission, which both provide high-resolution imagery under "complete, free and open"[42,43] licenses, will further increase data availability for monitoring forest management (for example, under the planned EU Observatory on changes in the world's forest cover)[44].

In summary, our results reveal a striking increase in forest harvesting in 26 European countries—a 49% increase in harvested forest area and a 69% increase in harvested biomass—for the years 2016–2018 compared with the average for 2011–2015, with potential implications for climate change mitigation from forest carbon sequestration and other ecosystem services. This type of timely and transparent monitoring of forest harvests is key for implementing more effective forest-based climate mitigation policies and for tracking the progress of country-based climate-mitigation targets. We contend that the carbon impact associated with increased forest harvesting in Europe, as observed in this study, will have to be counted towards post-2020 country-based EU climate targets[23,37]. We believe that the approaches we outline here for the monitoring of natural resources with big data will support future assessments of the potential trade-offs arising from the increasing demands on European forests from economic and ecological services. In addition, such approaches will improve the implementation of forest-related policies under the European Green Deal[14] and in meeting the greenhouse gas reporting and verification requirements under the Paris Agreement.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2438-y.

1. Forest Europe. *State of Europe's Forests 2015* (Ministerial Conference on the Protection of Forests in Europe, 2015); https://foresteurope.org/state-europes-forests-2015-report/.
2. Grassi, G. et al. The key role of forests in meeting climate targets requires science for credible mitigation. *Nat. Clim. Chang.* **7**, 220–226 (2017).
3. Masson-Delmotte, V. et al. (eds) *IPCC 2018: Global Warming of 1.5 °C* (Intergovernmental Panel on Climate Change, 2018); https://www.ipcc.ch/sr15.
4. Reid, W. V. et al. *Ecosystems and Human Well-being – Synthesis: A Report of the Millennium Ecosystem Assessment* (Island Press, 2005).
5. Duncker, P. et al. Classification of forest management approaches: a new conceptual framework and its applicability to European forestry. *Ecol. Soc.* **17**, 51 (2012).
6. Schelhaas, M.-J. et al. Actual European forest management by region, tree species and owner based on 714,000 re-measured trees in national forest inventories. *PLoS One* **13**, e0207151 (2018).
7. European Commission. *A New EU Forest Strategy: For Forests and the Forest-Based Sector Communication to the European Parliament Document no. 52013DC0659* (Publications Office of the European Union, 2013); https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52013DC0659.
8. Alkama, R. & Cescatti, A. Biophysical climate impacts of recent changes in global forest cover. *Science* **351**, 600–604 (2016).
9. Duveiller, G., Hooker, J. & Cescatti, A. The mark of vegetation change on Earth's surface energy balance. *Nat. Commun.* **9**, 679 (2018).
10. Luyssaert, S. et al. Trade-offs in using European forests to meet climate objectives. *Nature* **562**, 259–262 (2018); correction **567**, E13 (2019).
11. Le Quéré, C. et al. Global carbon budget 2017. *Earth Syst. Sci. Data* **10**, 405–448 (2018).
12. Shukla, P. R. et al. (eds) *IPCC, 2019: Climate Change and Land* (Intergovernmental Panel on Climate Change, 2019); https://www.ipcc.ch/srccl.
13. United Nations Framework Convention on Climate Change. *The Paris Agreement* (United Nations, 2015); https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement.

14. European Commission. *The European Green Deal Communication to the European Parliament Document no. 52019DC0640* (Publications Office of the European Union, 2019); https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2019%3A640%3AFIN.
15. Mandl, N. & Pinterits, M. (eds) *Annual European Union Greenhouse Gas Inventory 1990–2016 and Inventory Report 2018*. EAA Report no. 5/2018 (European Environment Agency, 2018); https://www.eea.europa.eu/publications/european-union-greenhouse-gas-inventory-2018.
16. Bastrup-Birk, A. et al. *European Forest Ecosystems – State and Trends*. EAA Report No. 5/2016 (European Environment Agency, 2016); https://www.eea.europa.eu/publications/european-forest-ecosystems.
17. Pilli, R., Grassi, G., Kurz, W. A., Viñas, R. A. & Guerrero, N. H. Modelling forest carbon stock changes as affected by harvest and natural disturbances. I. Comparison with countries' estimates for forest management. *Carbon Balance Manag*. **11**, 5 (2016).
18. Nabuurs, G.-J. et al. First signs of carbon sink saturation in European forest biomass. *Nat. Clim. Chang*. **3**, 792–796 (2013).
19. Kallio, A. M. I. & Solberg, B. On the reliability of international forest sector statistics: problems and needs for improvements. *Forests* **9**, 407 (2018).
20. Buongiorno, J. On the accuracy of international forest product statistics. *Forestry* **91**, 541–551 (2018).
21. Hansen, M. C. et al. High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853 (2013).
22. Gorelick, N. et al. Google Earth Engine: planetary-scale geospatial analysis for everyone. *Remote Sens. Environ*. **202**, 18–27 (2017).
23. Grassi, G., Pilli, R., House, J., Federici, S. & Kurz, W. A. Science-based approach for credible accounting of mitigation in managed forests. *Carbon Balance Manag*. **13**, 8 (2018).
24. Ericsson, K. & Werner, S. The introduction and expansion of biomass use in Swedish district heating systems. *Biomass Bioenergy* **94**, 57–65 (2016).
25. Bontemps, S. et al. Consistent global land cover maps for climate modeling communities: current achievements of the ESA's land cover CCI. In *ESA Living Planet Symp. 2013 CCI-4* (ESA, 2013).
26. Wolf-Crowther, M., Mozes, C. & Laczko, R. (eds) *Forestry in the EU and the World: A Statistical Portrait* (Eurostat, 2011); https://ec.europa.eu/eurostat/web/products-statistical-books/-/KS-31-11-137.
27. Santoro, M. *GlobBiomass – Global Datasets of Forest Biomass* https://pangaea.de/10.1594/PANGAEA.894711 (PANGAEA, 2018).
28. Noss, R. F. Assessing and monitoring forest biodiversity: a suggested framework and indicators. *For. Ecol. Manage*. **115**, 135–146 (1999).
29. Wilcox, B. A. & Murphy, D. D. Conservation strategy the effects of fragmentation on extinction. *Am. Nat*. **125**, 879–887 (1985).
30. FAOSTAT. *Forest Land* (Food and Agriculture Organization of the United Nations, 2019); http://www.fao.org/faostat/en/#data/GF.
31. Levers, C. et al. Drivers of forest harvesting intensity patterns in Europe. *For. Ecol. Manage*. **315**, 160–172 (2014).
32. Eurostat. *Economic Aggregates of Forestry* https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=for_eco_cp&lang=en (2019).
33. Forestry and Timber – UNECE. *Forest Products Annual Market Review, 2017–2018* (United Nations, 2018); https://www.unece.org/forests/fpamr2018.html.
34. European Commission. *Bioeconomy: the European way to use our natural resources* (European Commission, 2018); https://ec.europa.eu/knowledge4policy/node/33002_es.
35. European Parliament. *The Promotion of the Use of Energy from Renewable Sources Directive of the European Parliament Document no. 32018L2001* (Publications Office of the European Union, 2018); https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32018L2001.
36. Searchinger, T. D. et al. Europe's renewable energy directive poised to harm global forests. *Nat. Commun*. **9**, 3741 (2018).
37. European Parliament. *The Inclusion of Greenhouse Gas Emissions and Removals from Land Use, Land Use Change and Forestry in the 2030 Climate and Energy Framework Regulation of the European Parliament Document no. 32018R0841* (Publications Office of the European Union, 2018); https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2018.156.01.0001.01.ENG.
38. Achard, F. et al. Determination of tropical deforestation rates and related carbon losses from 1990 to 2010. *Glob. Change Biol*. **20**, 2540–2554 (2014).
39. Gibbs, H. K., Brown, S., Niles, J. O. & Foley, J. A. Monitoring and estimating tropical forest carbon stocks: making REDD a reality. *Environ. Res. Lett*. **2**, 045023 (2007).
40. Hansen, M. C., Stehman, S. V. & Potapov, P. V. Quantification of global gross forest cover loss. *Proc. Natl Acad. Sci. USA* **107**, 8650–8655 (2010).
41. Nabuurs, G.-J. et al. Next-generation information to support a sustainable course for European forests. *Nat. Sustain*. **2**, 815–818 (2019).
42. European Space Agency. *Copernicus Open Access Hub* https://scihub.copernicus.eu (2015).
43. Zhu, Z. et al. Benefits of the free and open Landsat data policy. *Remote Sens. Environ*. **224**, 382–385 (2019).
44. European Commission. *Stepping Up EU Action to Protect and Restore the World's Forests Communication to the European Parliament Document no. 52019DC0352* (European Commission, 2019); https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52019DC0352.

Methods

Forest mapping

In Europe, the characteristics of forests change considerably along climate gradients and among forest types. Consequently, there is not a common definition of a 'forest' but each country has adopted the definition that best fits national circumstances. The establishment of a national definition of a forest is essential to monitor changes in forest area and a prerequisite to develop a consistent monitoring system. The United Nations Framework Convention on Climate Change (UNFCCC) proposed that a 'forest' is an area of land of at least 0.05–1 ha and a minimum tree-crown cover of 10–30%, with trees that reach, or could reach, a minimum height of 2–5 m at maturity[45]. Inside these limits, EU countries selected their national forest definition for reporting purposes; EU regulation 2018/841 regarding land use, land use change and forestry[37] reports tabular values of the different tree-cover thresholds for each country. However, even small differences in forest definition might have amplified effects on amounts of biomass or stored carbon amongst others.

Forest cover and the relative changes were obtained combining data from the GFC maps[21] (which provide estimates on tree cover in the year 2000) with forest-area statistics from FAOSTAT. It should be noted that Hansen et al.[21] in their work refer to tree cover. As a consequence, a tree-cover threshold that defines forest cover must be selected to map forest cover from the GFC maps.

The Hansen maps of Global Forest Change. The Hansen maps of Global Forest Change[21] (GFC) version 1.6 are the results of a time-series analysis of the Landsat archive characterizing forest extent and forest change with a spatial resolution of about 30 m (the spatial resolution slightly varies along the latitude). The GFC maps consist of three layers: '2000 Tree Cover', 'Forest Loss Year' and 'Forest Cover Loss/Gain'. '2000 Tree Cover' is a global map of tree canopy cover (expressed in percentage) for the year 2000, where a 'tree' is defined as the canopy closure for all vegetation taller than 5 m in height. 'Forest Loss Year' refers to the year of gross forest-cover-loss event. Encoded as either 0 (no forest loss) or else a value in the range 1–18, representing forest loss detected primarily in the years 2001–2018, respectively. 'Forest Cover Loss/Gain' is defined as a stand-replacement disturbance or the complete removal of tree-cover canopy at the Landsat pixel scale. 'Gain' is defined as the inverse of loss, or a non-forest to forest change entirely within the period 2000–2012. Although forest-loss information is reported annually (in other words, there are annual maps for forest-loss disturbances), forest gain is reported as a 12-yr total, that is, it refers to the period 2000–2012 and is a unique layer that does not report the timing of the gain.

Our approach has limitations in the detection of small-scale silvicultural practices. Although the GFC clearly does not require full clear-cuts to detect forest-cover loss, it is not able to reliably capture partial removal of trees caused by forest thinning, selective logging, short cycle forestry (that is, less than 10 yr) or forest degradation when the tree-cover change is smaller than the Landsat spatial resolution. In addition, most changes occurring below the canopy cannot be detected by optical instruments, potentially leading further to an underestimation of actual harvest wood. It should also be noted that our analysis encompasses the 2004–2018 period, thus excluding the 2001–2003 period. The GFC dataset is based on the Landsat archive, and the temporal coverage throughout Europe for the first years is sparser, which can cause artefacts when calculating trends. Also, the GFC product is not fully consistent over the entire 2000-onward period. The ingestion of Landsat 8 from 2013 onwards leads to improved detection of global forest loss.

In terms of data acquisition, the analysis of Landsat images shows that the number of cloud-free images (defined as images with cloud cover less than 20%) over Europe gradually increases from 2013 to 2018 (Extended Data Fig. 9a). In particular, in the 2016–2018 period there is a 15% increase in Landsat image availability with respect to the preceding 3-yr period (2013–2015). In 2012, the number of images dropped substantially, owing to the decommission of Landsat 5.

However, our analysis shows that there is complete and frequent cloud-free land coverage of Landsat in Europe with more than seven cloud-free acquisitions per tile every year during the study period (2004–2018; Extended Data Fig. 9b). According to the authors of the GFC product, a minimum of seven acquisitions per year is sufficient to detect forest loss in Europe[46]. In fact, in temperate and boreal regions, forest recovery after harvesting (if occurring) is a much slower process compared to that occurring in tropical and subtropical regions, and the change in spectral signature persists for several months after the loss of vegetation and soil exposure. For these reasons we conclude that variation in image availability did not affect the results of our analysis, as the number of images collected was above the threshold required for a robust classification throughout the entire time series.

The only exceptions occurred in 2012, with longer satellite revisiting time in northern Europe, and in 2008 with a data gap in Fennoscandia, but this area presented marginal forest cover and forest loss throughout the whole study period.

FAOSTAT. The FAOSTAT Forestry database[30] provides annual production and trade statistics for forest products, primarily wood products such as roundwood, sawnwood, wood panels, pulp and paper. For many forest products, historical data are available from 1961. These statistics are provided by countries through an annual survey conducted by the FAO (Food and Agriculture Organization of the United Nations) Forestry Department. Within this study, we used 'Area of forest' data from FAOSTAT for each European country for the years 2000, 2005, 2010 and 2015[47].

From tree cover to forest cover

In this study, we present a simple approach towards defining for each EU26 country the minimum tree cover (percentage) that qualifies as forest using the GFC maps. For each country, we found that the tree-cover threshold needed to define a forest that minimizes the difference between national forest-area statistics from FAOSTAT and GFC estimates (Extended Data Fig. 1a). Specifically, we computed for 15 tree-cover classes—from 10% to 80% in 5% steps—the corresponding forest areas and selected the class that minimizes the difference between the national forest-area statistics collected in the FAOSTAT report for the year 2015 (hereafter FAOSTAT-2015); using the last published dataset is a common approach. To match the FAO definition of forest, we used a minimum mapping unit (MMU) of approximately 0.5 ha with a moving-window kernel. Specifically, in a square kernel 100 m × 100 m, we retain the forest only if there are more than 5 forest pixels in the GFC map, corresponding to about 0.45 ha. To explore the sensitivity of our analyses to the choice of tree cover, we replicated the analysis above using high and low tree-cover thresholds. A forest threshold sensitivity, S, (Extended Data Fig. 1b) was computed as

$$S = \frac{\text{Forest}_{\text{max}} - \text{Forest}_{\text{min}}}{\text{Forest}_{\text{rightThreshold}}} \times 100 \qquad (1)$$

Where $\text{Forest}_{\text{min}}$ represents the forest area obtained using a tree-cover threshold equal to 10%, $\text{Forest}_{\text{max}}$ represents the forest area obtained using a tree-cover threshold equal to 70% and $\text{Forest}_{\text{rightThreshold}}$ represents the forest area obtained using the correct tree cover (that is, the threshold that minimizes the difference with FAOSTAT-2015 estimates).

In other words, the forest threshold sensitivity represents how much the forest area would change by choosing strict or less strict thresholds (10% and 80% of tree cover) normalized by the actual forest area. If the forest sensitivity is, for instance, 120%, then using the two extreme thresholds for a forest definition (that is, tree cover equal to 10% and

70%) corresponds to forest areas that differ by 1.2 times the value of the actual forest area (as defined in Supplementary Fig. 1).

The results of this analysis show that national forest-areas change considerably according to the choice of the minimum tree-cover threshold and that this threshold varies by country, making it inappropriate to use a single threshold for the whole of Europe.

It should be noted that the GFC definition of forest is land-cover based, whereas the national forest inventories employ a land-use definition. For example, orchards are considered as forests in the GFC, whereas they are excluded from national forest inventories. Conversely, bare ground which has been affected by harvest operations is still called forest if it is expected to revert to forest by national forest inventories (land-use approach). Thus, the GFC maps can be used to produce a map of forest cover, with some caveats[48–50].

We note that the geographical extent of this study included 26 member states of the EU, including the UK, and excluding Cyprus and Malta—for which there are no data available from official government sources, or the forest coverage is scarce.

Comparing forest cover with different data streams

We compared our estimates of forest cover with estimates from the two existing datasets for EU26: FAOSTAT and LUCAS. FAOSTAT provides forest-area estimates for the years 2000 and 2010. LUCAS, the Land Use and Cover Area frame Survey carried out by Eurostat[51] (the statistical office of the EU), is an EU26-wide regular point-sample survey with a 2-km grid size that provides estimates for the years 2009, 2012 and 2015. Note that we used forest area from FAOSTAT for the year 2015 to define the tree-cover thresholds. However, a comparison using different years (see below, and Extended Data Fig. 2) gives further verification of our forest assessment.

To compare our calculated forest cover over the same years, we computed forest cover for the years 2000, 2009, 2010 and 2012 using the country-based tree-cover thresholds and considering a MMU of approximately 0.5 ha. We also took into account forest-gain information.

Extended Data Fig. 2 shows the comparison between FAOSTAT and GFC-derived forest area for 2000 and 2010. Note that for the GFC maps the temporal evolution of the forest area is always decreasing, whereas FAOSTAT often shows an increasing trend. This is probably because forest gain is difficult to capture with remote-sensing data. A decreasing trend in forest area for both GFC and FAOSTAT data is visible only for Finland and Portugal. The comparison shows a high level of agreement between the two datasets, which lends confidence to the assessment of remote-sensing-derived forest area.

The scatterplot analysis performed with FAOSTAT was also carried out with LUCAS data for 2009, 2012 and 2015, to have another independent source of information on forest area (Extended Data Fig. 2). The LUCAS data tend to provide larger estimates of forest area compared with GFC data. Such differences between forest estimates are probably due to the methodology: the LUCAS definition of forest is different from the FAO definition. Specifically, LUCAS uses a low tree-cover threshold—10%—and no MMU to define a forest (labelled as 'wooded area' in the dataset). In addition, changes in survey protocol for the 2009, 2012 and 2015 LUCAS campaigns might cause inconsistencies when datasets are compared over time.

Validation of the GFC maps with high-resolution imagery

We validated the GFC maps using high-resolution imagery from Google Earth. We performed two validation exercises aimed at testing the capability of the GFC for the detection of harvest patches of different sizes, designed as follows.

Validation exercise 1. We tested the GFC capabilities for forest-harvest patches of various sizes (hereafter, general validation). The purpose of this general validation was to assess the accuracy of the harvested area as derived from the GFC dataset (that is, the user accuracy). We did not attempt to quantify the omission errors. The general validation was carried out by analysing 620 patches of harvest with various size, randomly selected from seven countries (Poland, Ireland, France, Italy, Estonia, Sweden and Finland) for 2012 and 2017 to better sample the range of variability represented by different countries, climatic conditions, forest type and management system (620 patches in both 2012 and 2017). 26% and 37% of the patches for 2012 and 2017, respectively, could not be validated for lack of high-resolution imagery.

Validation exercise 2. This second validation effort was aimed specifically at testing our methods on big harvest patches (larger than 4.5 ha, hereafter the big-patch validation), as the increased occurrence of larger harvest areas is one of the main issues raised by this study. For the big-patch validation, we compared data from the same seven countries used in the general validation, and compared 2012 and 2017. For this exercise, forest patches consisted of at least 50 contiguous pixels (with a four-neighbours rule), that is, at least approximately 4.5 ha. We found 188 and 260 patches for 2012 and 2017, respectively.

For both the general and big-patch validations, samples were classified, using visual image interpretation, into four categories: 1) correct classification: the high-resolution images confirm the forest loss detected by the GFC maps in shape, position and timing (that is, the loss area in the high-resolution images is more than 50% of the loss area detected by GFC); 2) wrong classification: the forest loss detected by GFC is not visible in the high-resolution images; 3) partially correct (location and extent mismatch): the loss area in the high-resolution images is less than 50% of the loss area detected by GFC, mostly owing to image misregistration; and 4) partially correct (temporal mismatch): there is a temporal lag of maximum one year in the detection of GFC forest loss (generally, the actual loss happened the year before the loss reported by the GFC data).

Extended Data Fig. 3a reports the validation results by large (that is, ≥0.27 ha) and small (that is, <0.27 ha) forest-loss patches. It emerges that the classification capabilities are better in the year 2017 than in 2012, probably because Landsat 8 entered operation. As expected, the classification of small patches show a larger uncertainty (that is, the error in classification is 29% of cases instead of the 13% error observed for large patches in 2017). From these results we determine that, despite the larger uncertainty in the classification of small patches, the overall impact on our findings is limited, because patch sizes smaller than 0.27 ha represent less than 3% of the detected total harvested area in EU26 (Supplementary Fig. 1). The results of the big-patch validation clearly show that more than 84% of big forest patches (≥4.5 ha) are correctly classified and only 5% are wrongly classified (third row of Extended Data Fig. 3b). The remaining patches are either recorded with one year of delay (3%) or refer to harvest areas of different size (7%), owing to image misregistration.

This evidence confirms the robustness of our retrievals on the recent trend in harvest areas.

Spatial aggregation and major windstorm removal

To identify anomalies in forest management and to exclude extraordinary losses owing to natural disturbances that are not related to the normal management regime, we computed the annual percentage of forest loss at a 0.2° spatial resolution as the ratio between the area of forest loss during 2004–2018 and the area of forest cover in the year 2000, within each grid cell. Regions affected by forest fires, as detected by the European Forest Fire Information System (EFFIS) dataset, were masked out. EFFIS provides European Commission services and the European Parliament with updated information on wildland fires in Europe[52]. EFFIS provides shapefiles for European forest fires using remote-sensing imagery; specifically it maps burned areas by analysing daily images from MODIS at 250-m spatial resolution. Small burnt or unburnt areas below the spatial resolution of the MODIS imagery are not mapped; however, the area burned by fires detected by MODIS represents about 75% to 80% of the total area burned in the EU.

To generate Fig. 1 at the European scale, a common tree-cover threshold of 20% (instead of a country-specific threshold as used in the rest of the analysis) was used to define a 'forest'. We also excluded areas with sparse forest cover—that is, where forest cover in a gridcell of 0.2° is less than 10%. Aggregating to 0.2° also has another advantage, namely that this scale is simpler to map and visualize at the EU level, as shown in Figs. 1, 2b.

What is detected by satellites is a change in the percentage of forest cover that can either be attributed to forest management (that is, harvest) or disturbances (for example, pests, biotic disturbances and windstorms), and so we filtered out from our analysis areas affected by major windstorms. To do so, we assumed that major windstorms generally cause larger losses than the losses caused by forest management[53]. For each 0.2° grid cell we computed a threshold of the percentage of forest loss, which is calculated as:

$$\mathrm{Threshold}_{\mathrm{wind}} = \mathrm{median}(x) + 3 \times \mathrm{MAD}(x), \qquad (2)$$

where x is the time series of the percentage of forest loss from 2001 to 2018 and MAD is the median absolute deviation.

When the annual percentage of forest loss is greater than $\mathrm{Threshold}_{\mathrm{wind}}$, the the forest loss is attributed to windthrow. With this formula, we excluded major windstorms from our analysis. The resulting maps only remove major windstorms; forest loss from small and localized windstorms, pests and other diseases are not masked out. Note that $\mathrm{Threshold}_{\mathrm{wind}}$ was computed including the 2001–2003 period (later excluded from the analysis) to obtain more robust statistics.

Major windstorms are masked in Figs. 1, 2b. Patterns of major windstorms detected with our scheme show a good overlap with the tracks of major windstorms events in 2005, 2007 and 2009[53].

The major windstorms removal scheme has a major limitation, namely that short rotation forestry[54]—that is, areas characterized by intensive management—can be erroneously classified as major windstorms and thus excluded from our analysis. However, this limitation does not undermine the main findings of this study, as the rise in harvested forest area in the EU might be underestimated by excluding short-rotation forests.

A note of warning in Fig. 1 is warranted for Portugal, as during the period 2016–2018 the country experienced intense fires[52] that might have been only partially detected in our analysis (possibly owing to the limited spatial extent of individual events) and therefore erroneously considered as harvest area.

Land cover

The land cover data layer (at a resolution of 300 m) was obtained from the ESA GlobCover map[25] and harmonized to the 30-m^2 grid using a nearest-neighbourhood algorithm.

Patch size

We computed for each year and for each EU26 country the number of contiguous pixels—using a four-connected rule—of forest loss and its distribution. We excluded from the analysis regions affected by forest fires (using the EFFIS dataset) or major windstorms. For each year we computed the median of the number of connected pixels of forest loss. This median value is representative of the average patch size of harvested forest patches (Fig. 4). Combining country-level variations in harvested patch size and harvested forest area (as shown in Fig. 2a), it is possible to identify countries where the signs of the variation in harvested patch size and area are in opposition (that is, both patch size and area are either increasing or decreasing), as indicated by blue labels in Fig. 4, or not (red labels). Interestingly, in seven countries out of the 26, variations in the harvest area and patch size are in opposition. For example, in Sweden, the harvested area increased and the patch size decreased, although slightly (approximately 3%). Similarly, Austria, Bulgaria and Slovakia show an increase in the harvested forest area and at the same time a reduction in the patch size. This could suggest an increase in harvested forest area in smaller regions (for example, by private owners) or the application of less intensive management practices. Conversely, Belgium and Germany show an increase in the patch size and at the same time a reduction in the harvested area.

Silvicultural practice and harvest patch size

We conducted an analysis of changes in forest harvest size both at the European and also at the country level. We investigated the annual distribution of harvested forest area for five different classes of patch size, ranging from small patches (harvested forest area less than 0.27 ha) to large ones (harvested forest area greater than 7.2 ha) across all EU26 (Extended Data Fig. 5a) and at country level (Extended Data Fig. 5b). We note that the patterns for all EU26 and Finland are similar, with a major contribution from large patches of harvested forest. Conversely, Italy displays a dominance of harvested forest patches of size less than 3.6 ha, despite an increase in the number of big patches (>7.2 ha), which doubled from 2004 to 2016. These data provide information on the most common forest management practices applied at country level. On the one hand we have countries, such as Sweden, UK, Finland and Ireland, where larger harvested forest areas prevail, suggesting the application of clear-cut as the main management system. On the other hand, in Italy other silvicultural systems clearly prevail (such as the shelterwood system or a single-tree selection system): this is as a result of both the uneven age structure of the trees and to the smaller sizes of privately owned forests. It should be noted that, owing to calculation constraints, the sizes of the patches are calculated from the GFC map on a geographic coordinate system (that is, EPSG:4326) and not on an equal-area projection. As a consequence, slight errors in the area occur along latitude.

Harvested forest area at the country level and comparison with official harvest statistics

For each EU26 country, we compared the harvested forest area derived from the GFC maps and the amount of harvest volume removals reported by FAOSTAT. Harvest removals (that is, 'total roundwood production') are provided by FAOSTAT for each European country for the years 2004–2018, further corrected to account for possible inconsistencies, according to a previous analysis[17]. Harvest removals are expressed as volumes.

For this analysis, we excluded areas affected by forest fires (that is, the EFFIS archive), whereas areas affected by major windstorms were retained. In this way, we assume that storm-damaged timber is harvested and so that we are consistent with national harvest removal statistics that take into account salvage logging associated with wind damage; these generally exclude fires.

In Extended Data Fig. 6, the black line (normalized between zero and the maximum value of the harvested area for clarity) shows the harvest removals. Finally, the difference between Earth observation data and inventories is shown for the two countries with the largest forest sectors in the EU: Finland[55] and Sweden[56] (Supplementary Fig. 3), for which we have information on harvested forest area up to 2018 and 2016, respectively.

On the basis of this comparison between harvested forest area, official harvest removals (Extended Data Fig. 6) and National Forestry Action Programmes and other data sources (such as the National Forestry Accounting Plans (NFAP) recently published by the EU countries), we performed the following country-based analysis.

Austria. The GFC maps accurately reproduce the trend reported by harvest removals ($r = 0.65$; r, coefficient of correlation). This is also a result of the specific management system applied at national level, where the annual share of the final cut (the last of a series of cuts) of the total harvest is generally higher than 80% (NFAP Austria)[57]. These data series include both the amount of wood removed from salvage logging after major windstorms (in 2007–2008)[58], and also the area affected by those disturbance events.

Belgium. Uncertainties in official harvest removal data, and the peak in 2010—probably due to a windstorm[59]—reported only by the GFC, may explain the lack of correlation between the two time series.

Bulgaria. The high uncertainty in official harvest removal data[60], the effects of unregistered logging and heterogeneous silvicultural systems applied at country level (including simple coppices and coppices in conversion to high forests) may explain the low correlation between GFC and harvest removals.

Croatia. The poor correlation with the GFC maps is probably due to the specific forest management systems applied at national level, including the shelterwood system (largely applied to broadleaves), and the selective cut system (applied to unevenly aged forests, which cover about 20% of the total forest area). Moreover, silvicultural treatments are still partially influenced by ongoing demining activities, owing to the war that involved Croatia during the 1990s (NFAP Croatia)[61].

Czech Republic. The GFC maps represent fairly well the amount of harvest provided by final cut (on average, 43% of the total removals) and, partially owing to salvage logging, equal to about 41% of the total removals during the last decade (NFAP Czechia)[62]. The peak of harvest as reported by both these time series since 2016 is probably the result of salvage logging, as a consequence of windstorms and bark beetle attacks that have occurred during recent years[63].

Denmark. The lack of correlation with the GFC data is due to both some uncertainty in the estimates reported by harvest removal data (generally underestimated before 2014), and also to the increasing amount of primary residues removed from forests from 2011 onward (NFAP Denmark)[64]. Owing to this activity, recent harvest removal data also include wood used for energy, mainly provided by branches and other wood materials.

Estonia. Data from the GFC are consistent with harvest removals, and probably include the amount of area affected both by final cut and also by salvage logging after major disturbance events.

Finland. Harvest data reported by official statistics is well correlated with data from the GFC ($r = 0.56$). Taking into account the information reported by the 2018 statistical yearbook for forestry in Finland[55] (Supplementary Fig. 3), we can infer that the GFC can be compared to the area affected by clear-cut (about 135 kha yr^{-1} for the period 2001–2016) and final removals within the shelterwood system (about 43 kha yr^{-1} for the period 2001–2016). Both these data series, however, are only partially correlated with the annual amount of harvest removed at country level ($r = 0.53$). This is probably due to the following: (1) the harvest from thinnings is not negligible, because thinning represents about 66% of the total area affected by fellings at country level (average of the period 2004 and 2015); and (2) the different biomass density per unit of area between the northern and southern part of the country certainly reduces the correlation between the two variables. Nevertheless, the increasing amount of harvest detected by the GFC during recent years was recently confirmed by the data reported by the National Resource Institute of Finland[65], highlighting that in 2018, a total of 78.2 million cubic meters of roundwood was harvested from Finnish forests, 8% more than in the previous year and, compared with the average of the preceding ten-year period, amounted to an increase of nearly 25%.

France. The GFC represents fairly well the amount of harvest from final cut and salvage logging after major natural disturbances (indeed, they clearly highlight the effect of the windstorm that occurred in 2009, which explains the peak of harvest removals reported for 2010). Owing to the complex structure and heterogeneity of the management systems applied in France (including coppice with standards, and mixed forests where coppices and high forests coexist in the same area), and also the difficulty in determining the different biomass densities per unit of area, the GFC can probably detect only part of the silvicultural treatments and of the overall harvest applied at the country level ($r = 0.33$).

Germany. Harvest data reported by official statistics is well correlated with data from the GFC ($r = 0.56$), and can be compared with the amount of harvest from final cut and salvage logging after major natural disturbances (the data clearly highlight the effect of the windstorms that occurred in 2007 and 2010).

Greece. Harvest data reported by official statistics is partially correlated with data from the GFC ($r = 0.42$). This is due both to the high uncertainty of harvest statistics[19,20] and also to the specific characteristics of this country, which is mainly covered by unevenly aged forests that are generally treated with selective cut systems.

Hungary. The GFC maps do not reproduce the pattern in the forest harvest data, probably because it cannot reproduce the sharp increase in total forest area as reported in official statistics[66] (between 2000 and 2015 the total forest area grew by 8%, from 1,908 kha in 2000 to 2,069 kha in 2015)[1].

Ireland. As for Hungary, the GFC does not reproduce the trend in the harvest data, probably because it cannot reproduce the sharp increase in total forest area as reported in official statistics (between 2000 and 2015 the total forest area grew by 19%, from 635 kha in 2000 to 754 kha in 2015)[1].

Italy. Owing to the high uncertainty of official harvest removal data[19,20] and to the specific characteristics of this country (unevenly aged forests cover about 30% of the total forest area)[67], and because the biomass density may vary within the country owing to different climatic conditions, the GFC can only partially reproduce the trend reported by harvest statistics.

Latvia and Lithuania. Even if the average share of harvest provided by clear cut is equal to about 70–80%[17], the GFC can only partially reproduce the trend reported by harvest removal data. This is specifically due to the decreasing amount of area affected by harvest detected by GFC maps between 2012 and 2015. For both these countries, even if the absolute amount of harvest has been generally increasing since 2010, the relative share of final cut to thinnings decreased, at least for some species (NFAP Lithuania)[68].

Luxembourg. The GFC maps can reasonably reproduce the trend reported by harvest removal data ($r = 0.60$). This is due to the specific management system applied at national level, where the annual share of the final cut is generally higher than 90%[17].

The Netherlands. The lack of statistical correlation between official harvest removals and GFC data may be due to different reasons. The data for harvest removals was extremely homogeneous in time until 2013, when, owing to an abrupt increase of coniferous wood removals, the total amount of harvest increased by about 16%[69]. Conversely, GFC data shows a peak in 2010 (when removals increased by 6% compared to 2009), and no major variation is reported after 2013. Neither the

GFC nor harvest removal data highlight any substantial deviation in 2007 when about 0.25 million m^3 in biomass volume was damaged by windstorms.

Poland. Overall, the GFC can reproduce the trend reported by harvest removal data ($r = 0.62$), at least for the quota compared to the amount of harvest provided by clear cut, equal to about 48% of average annual removals reported by the country since 2004 (NFAP Poland)[70].

Portugal. Despite the very heterogeneous silvicultural systems applied at country level (including unevenly aged forests), the GFC is well correlated with official harvest removal data ($r = 0.75$). This is probably also due to the relatively high proportion of *Eucalyptus* plantations among all forest area, many of which are managed through clear cuts.

Romania. Large uncertainties in official harvest statistics[19,20], unregistered logging and the various silvicultural treatments applied at country level (including unevenly aged forest systems) considerably reduce the correlation between GFC data and the official harvest removal data ($r = 0.39$).

Slovakia and Slovenia. The GFC data can adequately reproduce the trend reported by harvest removal data ($r = 0.73$ for both these countries). This is also due to the specific management system applied at a national level, which is largely based on clear cut (for Slovakia, the annual share of harvest provided by the final cut is generally higher than 70%)[71].

Spain. Owing to the specific characteristics of this country, which is largely covered by unevenly aged forests that are managed through a single tree selection system, the GFC maps can only partially reproduce ($r = 0.44$) the trend reported by harvest removal data[72].

Sweden. The lack of correlation between the GFC data and harvest-removal data is probably due to: (1) when large disturbance events occurred, salvage logging (for sanitary reasons) had the priority on clear cut, the area of which was indirectly reduced (for this reason, probably, the GFC does not highlight the effect of the two windstorms that occurred in 2005 and 2007); (2) remote-sensing estimates and harvest statistics at the country scale may not show a statistical correlation because the biomass density per unit of area differs greatly over the country in space (that is, between the northern or southern part of Sweden); and (3) for this country, final felling covered (in terms of area) about 37% of the area annually affected by fellings between 2000 and 2015[56]. This area is not statistically correlated with the total amount of wood removed during the same period, as reported by the same data source ($r = 0.48$). Despite that, official statistics on the notified area (larger than 0.5 ha), affected by final felling are consistent with the GFC (see Supplementary Fig. 3) and highlight that the size of this area increased by 13% in 2018 in comparison with the previous year, and compared with the average of the period 2011–2015, increased by nearly 17%. Considering that these statistics only report the "notified area larger than 0.5 ha"[56], whereas the GFC probably includes a broader share of management practices, we can infer that in Sweden the GFC maps adequately represent the variation in the relative amount of area affected by final felling.

United Kingdom. Overall, the GFC maps can reproduce the trend reported by harvest removal data ($r = 0.44$). Some peaks reported by the GFC in 2012 could be due to the indirect effect of exceptional fires that were not properly filtered out by the preliminary analysis performed on these disturbances[73].

Inconsistency between remote-sensing-based estimates (that is, the harvested area) and national statistics on harvest removals may be due to the specific silvicultural practices of the country and to the accuracy and time resolution of official harvest statistics. Concerning specific silvicultural practices, owing to the spatial scale of the GFC dataset, the detected harvested area is limited to management schemes that lead to the complete removal of trees on a minimum spatial scale of 30 m. Small-scale silvicultural practices such as thinning or selective logging—which are relevant in some EU countries—could therefore not be fully detected. The second aspect refers to the limitation of official statistics, which in some countries may be suboptimal because they are infrequently updated or are incomplete owing to unregistered or illegal logging. In these cases, the use of independent remote-sensing data, such as that provided by this study, could help in improving and act as a complement to national statistics.

We also performed a country-based assessment on the impact of thinning and selective logging on the total harvest (Supplementary Table 1). In this analysis, we reported the share of final cut for the managed area or, in the case of the Carbon Budget Model[17], volume from the evenly aged forests. National statistics highlight how thinnings or selective logging (on evenly and unevenly aged forests, respectively) is relevant only for a few EU countries (for example, Italy, France, and Croatia, as indicated in the previous sections). Also, low values of the share of clear cut (for example, as in Italy) may not hamper GFC statistics, because they partially include forest thinnings and other silvicultural practices such as salvage logging.

Potential drivers of change in harvested-forest area

Increasing harvest demand, as detected by our study, is potentially due to the combined effect of endogenous and exogenous drivers.

Endogenous drivers. are those deriving from forest characteristics (such as age–class distribution) that may affect the amount and temporal dynamic of the wood available for harvest even under a constant management system.

Exogenous drivers. include on one hand natural disturbances such as forest fires, heavy snow load and windthrow (which affect both the age structure and management practices), and, on the other hand, political, social or economic factors that lead to a modification of management practices applied with respect to a reference period, for example, to satisfy an increasing wood demand.

Quantifying and disaggregating the impact of the single drivers is challenging. Taking into account the effect of ageing and assuming the continuation of current management practices applied by the 26 studied EU countries between 2000–2009, it is estimated that, at the EU level, harvest volumes are expected to increase by 9% in the period 2021–2030 relative to the period 2000–2009[23]. Assuming a gradual increase in the harvest owing to ageing we should therefore expect a 0.45% increase per year. Similarly, another work[74] foresees a sustainable increase in harvest of 19%, owing to ageing, for the period 2009–2050 (equivalent to 0.46% per year).

Considering that the increase observed with satellite records occurred in the latter half of the decade (2016–2018), we estimate that over this timespan a maximum increase of about 4% by volume could be ascribed to forest ageing, which corresponds to about 8% of the observed increase in the harvested biomass. From this we can infer that endogenous drivers, as defined above, have had only a minor role in the recent sharp increase in harvest and that exogenous factors dominated.

Among exogenous drivers, the expansion of activities on the basis of demand for wood products (economic drivers) might have affected the forest sector, as reported in official statistics from UNECE and FAO[33] and Eurostat[75]. In fact, forest harvest is unlikely to increase when there is no rise in market demand for wood products. In northern and central–eastern Europe, where the relative contribution of the forest sector to GDP is the largest (2.1% and 1.3%, respectively, in 2010)[1], the higher demand from sawmills during the last years was probably one of the major drivers of the increasing timber harvest[33]. For example, in Croatia sawn-hardwood production grew by 89% in the five years to 2017, and in the Czech Republic and Slovakia particleboard production grew by 10% and 6.5%, respectively, in 2017 compared with the previous year[33]. In addition, fuelwood removals increased at the EU26 level from around 70 $\mathrm{Mm^3}$ to about 99 $\mathrm{Mm^3}$ (+41%) between 2000 and 2015[76]. UNECE[33] also confirms a substantial increase of EU harvest in 2013–2017 compared to 2007, with three countries standing out: Poland (+19.5%), Finland (+12.2%) and Sweden (+7.5%).

International trade, sometimes linked to political factors, may also affect the harvest demand at the national level. This was, for example, the case in some north European countries (such as Finland and Estonia), where, since 2009, the collapse of exports of roundwood from Russia indirectly affected internal harvest demand. Conversely, in some central European countries (such as the Czech Republic, Hungary and Slovenia), exports have strongly increased since 2014, encouraged not only by increasing roundwood demand coming from Germany (where imports increased by 30% since 2014), but also by from other EU26 countries (such as the UK and Croatia), and more recently, from China.

Concerning the increase in wood demand and its market, in the EU the application of the 'Energy from Renewable Sources' directive[35] and the bio-economy strategy[34] (started in 2012) are setting binding targets and increasing wood demand for bioenergy needs, with an established target of at least 32% renewable energy by the year 2030[34]. Specifically, the EU renewable energy directive[35] raised concerns about increasing harvested wood for bioenergy use[36]. In the ongoing shift from coal to biomass, wood is currently responsible for more than 60% of the renewable energy supply in Europe[75].

The outputs of forestry and connected secondary activities (Extended Data Fig. 7) increased by 13% in EU28 from 2012 to 2016[32], whereas in countries that show the largest increases in harvest—such as Poland, Portugal, Romania, Slovenia, Finland and Sweden—the rise was almost twofold (even if, for all these countries, statistics refer to the period 2008–2016).

The percentage of change in harvest area from 2008 to 2016 (or from 2012 when 2008 data is not available) as retrieved from remote sensing and from forestry market statistics are reported in brackets in the labels of Extended Data Fig. 7. Note that the quality of the Eurostat data varies from country to country, and some outliers (for example, France in 2014) seem questionable. Both UNECE and Eurostat indicators on wood products are heavily influenced by many other factors that can independently affect the true amount of forest harvest. However, these statistics give an overall indication of existing trends and potential drivers.

Concerning the potential effects of policy changes, the key role of the forest sector within the bioeconomy market has been supported by specific political initiatives in several EU countries. For example, this is the case in Slovenia, where specific financial incentives have actively supported the forest sector during recent years[77]. By contrast, in Sweden[78], as in other north European countries where production subsidies were abolished, the increase in felling during recent years is probably due to the increasing demand for forest raw materials by the forest industry.

A relevant recent element in the policy context is the EU regulation for the Land Use, Land Use Change and Forestry sector in the EU 2030 climate target[37], which aims to improve the assessment of the carbon impact of additional actions in "managed forest land"[23]. This regulation sets forest reference levels: country-based estimates of greenhouse gas emissions and removals in managed forest lands. The regulation has been strongly debated in scientific and policy contexts, and sometimes perceived as a possible limitation on potential future increases in harvest[79,80]. This knowledge might have triggered a more rapid increase in forest harvest in some countries, compared with what would have otherwise occurred. However, we could not find any direct evidence that this EU regulation is a reason for the increase in harvest.

A final set of exogenous drivers that may have affected forest-harvest intensity include natural disturbances such as windstorms, heavy snow load, forest fires and pest outbreaks. If the medium-term trend is mainly controlled by economic, political and legislative factors, salvage logging can represent the main driver affecting year-to-year fluctuations in total harvest at the country, regional or even at the EU level. As highlighted in Extended Data Fig. 7, this was the case in Austria (2007–2008), Czech Republic (2016–2018), France (2009–2010), Finland (2017–2018), Germany (2007 and 2010), Slovakia (2005), Slovenia (2014) and Sweden (2005 and 2007). Estimating the effect of natural disturbances on harvest statistics is challenging, because a fraction of the biomass will be directly removed through salvage logging, and the remaining will be harvested during the following years through normal silvicultural practices such as thinnings and clear-cuts. Despite this uncertainty, it is important to notice that at the EU level the amount of harvest owing to salvage of storm residue is somewhat limited. For instance, in the period 2000–2012 forest harvest owing to storms was on average equal to 13 $\mathrm{Mm^3\,yr^{-1}}$—that is, about 2.7% of the average total amount of harvest removed within the same period[17]. These events can generate large spatial and inter-annual variability so that at the country scale and for selected years the importance of salvage logging can be very relevant. For example, for the Czech Republic, the share of harvest provided by salvage logging in 2007 and 2017 was equal to about 83% and 60%, respectively. However, during the recent years characterized by an abrupt increase of harvest rate, there have been no major windthrow events at the European scale that may have contributed substantially to the observed trend. Moreover, as highlighted above, generally there is a mutual relation between salvage logging (for sanitary reasons) and ordinary management practices (such as clear cut) the affected area of which is indirectly reduced when large disturbance events occur.

Summarizing these considerations, we can conclude that the largest share (up to 90%) of the increasing amount of harvest detected during recent years is most probably due to exogenous drivers, whereas about 10% was the result of forest ageing. At the European scale, natural disturbances (which have probably affected inter-annual variations and trends) have been factored out from the analysis. Ultimately, recent changes in socio-economic and political contexts are thus the most probable driver of the observed patterns.

Above-ground biomass analysis

AGB values for harvested forest were obtained from ESA GlobBiomass, a global dataset of forest biomass at a resolution of 100 m for the year 2010[27]. Specifically, the AGB analysis quantifies the mass of all living trees excluding stump and roots, expressed as the oven-dry weight of the woody parts (stem, bark, branches and twigs) in units of $\mathrm{Mg\,ha^{-1}}$. The AGB estimates were obtained from space-borne synthetic-aperture radar (ALOS PALSAR, Envisat ASAR), optical (Landsat 7), lidar (ICESat) and auxiliary datasets with multiple estimation procedures[27]. The AGB map was resampled at the spatial resolution of the GFC (that is, resampled from 100 m to 30 m) and (as the AGB map refers to 2010) to update it to the year of forest loss from 2011 onwards, we assigned an AGB value of zero to those pixels with forest loss, meaning that forest loss was considered as a total AGB loss.

Forest biomass growth was retrieved from ref. [1]. The average biomass growth rate (Gr, expressed as an annual percentage) has been computed for five geographical regions in Europe (north, central west, central east, south west and sout east; see Extended Data Fig. 10a) as

$$\mathrm{Gr} = \frac{\mathrm{Gs}_{2015} + F_{2010-2015} - \mathrm{Gs}_{2010}}{\mathrm{Gs}_{2010}} \times \frac{1}{\mathrm{Ys}} \qquad (3)$$

Gs_{2010} and Gs_{2015} are the total growing stock in 2010 and 2015, respectively, Ys is the number of years between 2010 and 2015 (five) and $F_{2010-2015}$ is the total amount of fellings removed within the same period. We converted relative into absolute biomass growth rates (from percentage to t ha^{-1} yr^{-1}) on the basis of the AGB map and forest-area estimates by country from the GlobBiomass[27] and GFC[21] datasets, as shown in Extended Data Fig. 10b. As expected, the results show that absolute growth rates are higher in the temperate forests of central Europe and lower in boreal and Mediterranean regions.

Again, regions affected by forest fires (from EFFIS data) and major windstorms were excluded from our analysis. We note that resampling the biomass data from 100 m to 30 m is an approximation that introduces uncertainty in the biomass-loss estimates.

The analysis of AGB loss was carried out at the European and country level. Extended Data Fig. 8 shows the percentage of AGB harvested per year in a 0.2° grid cell, Supplementary Fig. 2 shows the pixel-wise R^2 regression between harvested forest area and biomass, and Supplementary Fig. 4 shows the percentage national contribution of the European harvested forest biomass during 2016–2018.

As expected, the pixel-wise correlation between harvested forest area and harvested forest biomass is high over the spatial domain (Supplementary Fig. 2), because harvested forest area and biomass are closely linked.

Supplementary Figs. 5 and 6 show the average harvested area for five biomass-density classes for the period 2011–2015 (left) and 2016–2018 (right) for Finland, Sweden, Poland and Italy. The patterns of Supplementary Fig. 5a show that the contribution of evergreen forests in the AGB range 50–150 t ha^{-1} dominate, whereas the contribution from forests with very high AGB (that is, greater than 150 t ha^{-1}) is negligible. Sweden (Supplementary Fig. 5b) shows patterns that are similar to those in Finland, although the quota of harvested biomass greater than 200 t ha^{-1} is higher. Conversely, Poland (Supplementary Fig. 6a) exhibits a dominance of mixed forests in the range 100–200 t ha^{-1}, indicating a different distribution of forest age and structure.

Cloud-computing platform: Google Earth Engine

Google Earth Engine is a cloud-based infrastructure that enables "access to high-performance computing resources for processing very large geospatial datasets"[22]. It consists of "a multi-petabyte analysis-ready data catalogue co-located with a high-performance, intrinsically parallel computation service"[22]. The data catalogue hosts a large repository of publicly available geospatial datasets, including the Landsat archive, the GFC maps[21], and land-cover, topographic and socio-economic datasets. From 2015, the Copernicus Sentinel sensor data are also included. The catalogue is accessed and controlled through an Internet-accessible application programming interface (API) that enables prototyping and visualization of results.

All data extraction for this study was performed in Google Earth Engine, which provides the ability to compute pixel-level or country-based statistics and analyse the entire data records of the GFC maps as well as ancillary land cover data with high computational efficiency, and without the need to retrieve and download huge amounts of data.

Data availability

To ensure full reproducibility and transparency of our research, we provide all of the data analysed during the current study. The data are permanently and publicly available on a Zenodo repository, https://doi.org/10.5281/zenodo.3687090.

Code availability

To ensure full reproducibility and transparency of our research, we provide all of the scripts used in our analysis. Codes used for this study (Google Earth Engine and R scripts, the harvest-removals dataset and shapefiles of the validation) are permanently and publicly available on a Zenodo repository, https://doi.org/10.5281/zenodo.3687096.

45. UNFCCC. *Report of the Conference of the Parties on its Seventh Session* (United Nations Framework Convention on Climate Change, 2001); http://unfccc.int/resource/docs/cop7/13a01.pdf.
46. Potapov, P. V. et al. Eastern Europe's forest cover dynamics from 1985 to 2012 quantified from the full Landsat archive. *Remote Sens. Environ*. **159**, 28–43 (2015).
47. Food and Agriculture Organization of the United Nations. *Global Forest Resources Assessment 2015* (Food and Agriculture Organization of the United Nations, 2015); http://www.fao.org/forest-resources-assessment/en/.
48. Tyukavina, A. et al. Types and rates of forest disturbance in Brazilian Legal Amazon, 2000–2013. *Sci. Adv*. **3**, e1601047 (2017).
49. McNicol, I. M., Ryan, C. M. & Mitchard, E. T. A. Carbon losses from deforestation and widespread degradation offset by extensive growth in African woodlands. *Nat. Commun*. **9**, 3045 (2018).
50. Baccini, A., Walker, W., Carvalho, L., Farina, M. & Houghton, R. A. Response to Comment on "Tropical forests are a net carbon source based on aboveground measurements of gain and loss". *Science* **363**, eaat1205 (2019).
51. Ballin, M., Barcaroli, G., Masselli, M. & Scarnó, M. *Redesign Sample for Land Use/Cover Area Frame Survey (LUCAS) 2018* (Eurostat, 2018).
52. EFFIS. *Statistics Portal* https://effis.jrc.ec.europa.eu/static/effis.statistics.portal (EFFIS, 2019).
53. Gardiner, B. et al. *Destructive Storms in European Forests: Past and Forthcoming Impacts* (European Forest Institute, 2010).
54. Hengeveld, G. et al. A Forest Management Map of European Forests. *Ecol. Soc*. **17**, 53 (2012).
55. Vaahtera, E. et al. *Finnish Forest Statistics* (Natural Resources Institute Finland (Luonnonvarakeskuksen), 2018); http://jukuri.luke.fi/handle/10024/543098.
56. *Sweden Statistical Database* http://pxweb.skogsstyrelsen.se/pxweb/en/Skogsstyrelsens%20statistikdatabas/?rxid=c8c5a437-bd93-4972-8060-50ea3026db1b (2019).
57. *National Forestry Accounting Plan for Austria* (Bundesministerium Nachhaltigkeit und Tourismus, 2018); https://www.fern.org/fileadmin/uploads/fern/Documents/NFAP_Austria.pdf.
58. Foglar-Deinhardstein, A., Hangler, J. & Prem, J. (eds) *Sustainable Forest Management in Austria, Austrian Forest Report 2008* (Federal Ministry of Agriculture, Republic of Austria, 2008).
59. Institut Européen De La Forêt Cultivée. *Forestorm* http://www.iefc.net/storm/.
60. Stoeva, L., Markoff, I. & Zhiyanski, M. *National Forestry Accounting Plan for Bulgaria, Including Forest Reference Levels for the Period 2021–2025* (Ministry of Environment and Water, Republic of Bulgaria, 2018); https://www.moew.government.bg/static/media/ups/articles/attachments/NFAP_final_ENe30a12d4d7dff0f589ad84b69bb90b12.pdf.
61. *National Forestry Accounting Plan For The Republic Of Croatia* (Ministry of Environment and Energy and Ministry of Agriculture, Republic of Croatia, 2018); https://mzoe.gov.hr/UserDocsImages//KLIMA/SZKAIZOS//NFAP_Croatia.pdf.
62. Ministry of Agriculture & Ministry of the Environment, Czech Republic. *National Forestry Accounting Plan For The Czech Republic, Including A Proposed Forest Reference Level* (Ministry of Agriculture and Ministry of the Environment, Czech Republic, 2018); https://www.fern.org/fileadmin/uploads/fern/Documents/NFAP_Czech_Republic.pdf.
63. Cienciala, E. et al. Recent spruce decline with biotic pathogen infestation as a result of interacting climate, deposition and soil variables. *Eur. J. For. Res*. **136**, 307–317 (2017).
64. Johannsen, V. K., Nord-Larsen, T., Scott Bentsen, N. & Vesterdal, L. *Danish National Forest Accounting Plan 2021–2030* (University of Copenhagen, 2019); https://efkm.dk/media/12542/danishnationalforestaccountingplan_2019.pdf.
65. *Finnish Forest Sector Economic Outlook* https://www.luke.fi/en/finnish-forest-sector-economic-outlook/ (2019).
66. Somogyi, Z. et al. *National Forest Accounting Plan of Hungary (2021–2025)* (2018); https://cdr.eionet.europa.eu/hu/eu/mmr/lulucf/envxbyrxa/.
67. Gasparini, P., Di Cosmo, L., Cenni, E., Pompei, E. & Ferretti, M. Towards the harmonization between National Forest Inventory and Forest Condition Monitoring. Consistency of plot allocation and effect of tree selection methods on sample statistics in Italy. *Environ. Monit. Assess*. **185**, 6155–6171 (2013).
68. State Forest Service, Environment Protection and Forest Department & Aleksandras Stulginskis University. *National Forestry Accounting Plan by Lithuania* (Ministry of the Environment, 2018); https://www.fern.org/fileadmin/uploads/fern/Documents/NFAP_Lituania.pdf.
69. FAOSTAT. *Forestry Production and Trade* http://www.fao.org/faostat/en/#data/FO (Food and Agriculture Organization of the United Nations, 2019).
70. Ministry of the Environment. *Poland National Forestry Accounting Plan (NFAP)* (2018); https://bip.mos.gov.pl/fileadmin/user_upload/bip/strategie_plany_programy/Krajowy_Plan_Rozliczen_dla_Lesnictwa/NFAP_2018_POLAND_ENG_FINAL.pdf.
71. Barka, I., Priwitzer, T. & Hušťáková, E. *National Forestry Accounting Plan of the Slovak Republic Including a Proposed Forest Reference Level* (Ministry of Agriculture and Rural Development of the Slovak Republic, National Forest Centre, 2018); http://www.mpsr.sk/out.php?urlid=77.
72. *National Forestry Accounting Plan for Spain* (2018); https://www.miteco.gob.es/es/cambio-climatico/temas/mitigacion-politicas-y-medidas/nfap_env20_tcm30-506250.pdf.
73. United Nations Climate Change. *National Inventory Submissions 2019* https://unfccc.int/process-and-meetings/transparency-and-reporting/reporting-and-review-under-the-convention/greenhouse-gas-inventories-annex-i-parties/national-inventory-submissions-2019 (UNFCCC, 2019).
74. Nabuurs, G.-J., Arets, E. J. M. M. & Schelhaas, M.-J. Understanding the implications of the EU-LULUCF regulation for the wood supply from EU forests to the EU. *Carbon Balance Manag*. **13**, 18 (2018).
75. Joint Research Centre. *Brief on Biomass for Energy in the European Union* (European Commission, 2019); http://publications.jrc.ec.europa.eu/repository/bitstream/JRC109354/biomass_4_energy_brief_online_1.pdf.

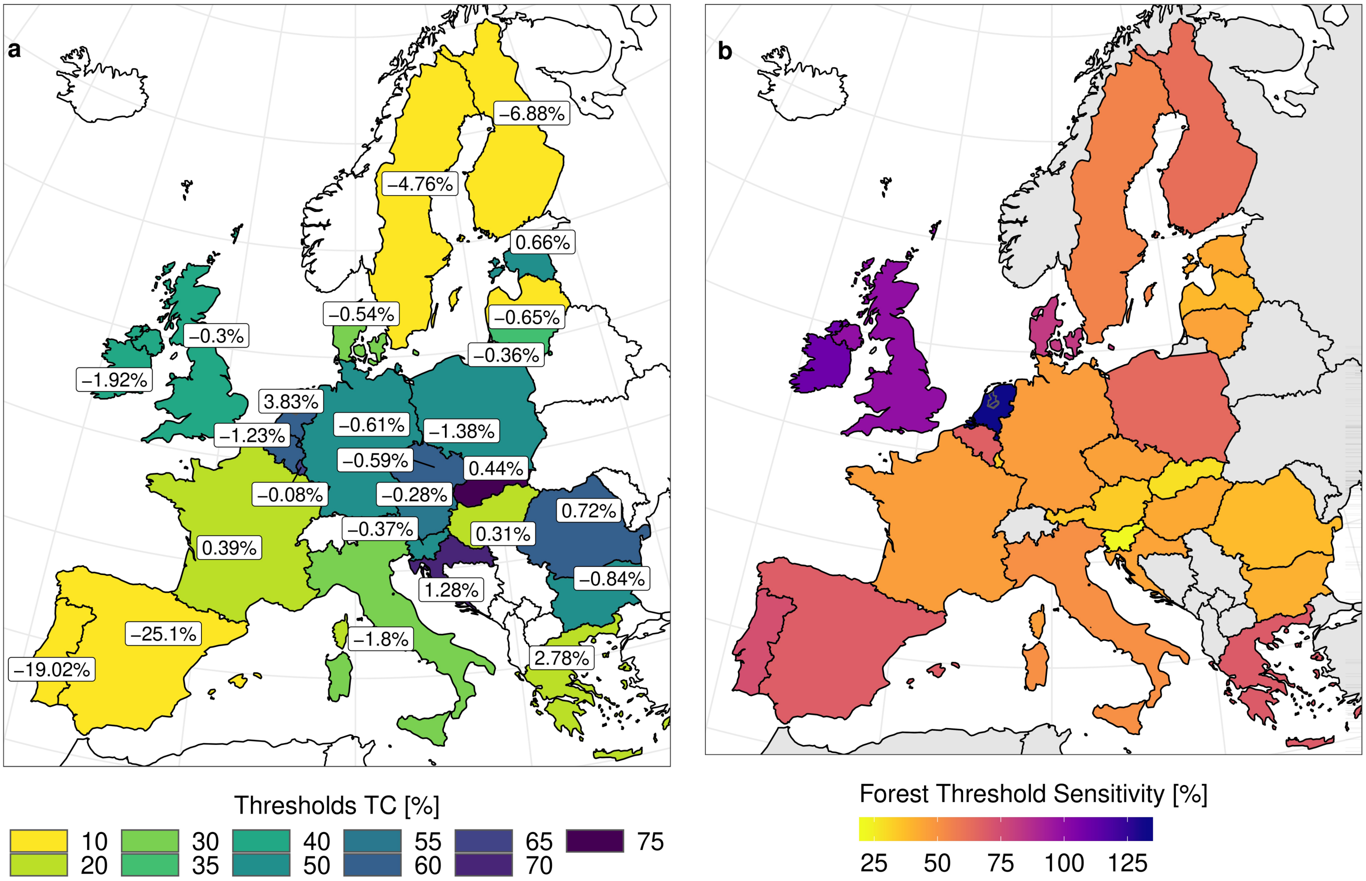
76. Hoepffner, N. et al. *Biomass Production, Supply, Uses and Flows in the European Union: First Results From an Integrated Assessment* (European Commission, 2018); https://publications.europa.eu/en/publication-detail/-/publication/358c6d4b-1783-11e8-9253-01aa75ed71a1/language-en.
77. *Slovenian National Forestry Accounting Plan, Including Forest Reference Levels – Draft* (2018); https://www.fern.org/fileadmin/uploads/fern/Documents/NFAP_Slovenia_translated.pdf.
78. *National Forestry Accounting Plan for Sweden* (2018); https://www.regeringen.se/495799/contentassets/38eaf6f23f284fb0a440b0742fe7bcf7/national-forestry-accounting-plan-for-sweden.pdf.
79. Grassi, G. et al. Wrong premises mislead the conclusions by Kallio et al. on forest reference levels in the EU. *For. Policy Econ*. **95**, 10–12 (2018).
80. Kallio, A. M. I., Solberg, B., Käär, L. & Päivinen, R. Economic impacts of setting reference levels for the forest carbon sinks in the EU on the European forest sector. *For. Policy Econ*. **92**, 193–201 (2018).

Acknowledgements The views expressed are purely those of the writers and may not in any circumstances be regarded as stating an official position of the European Commission.

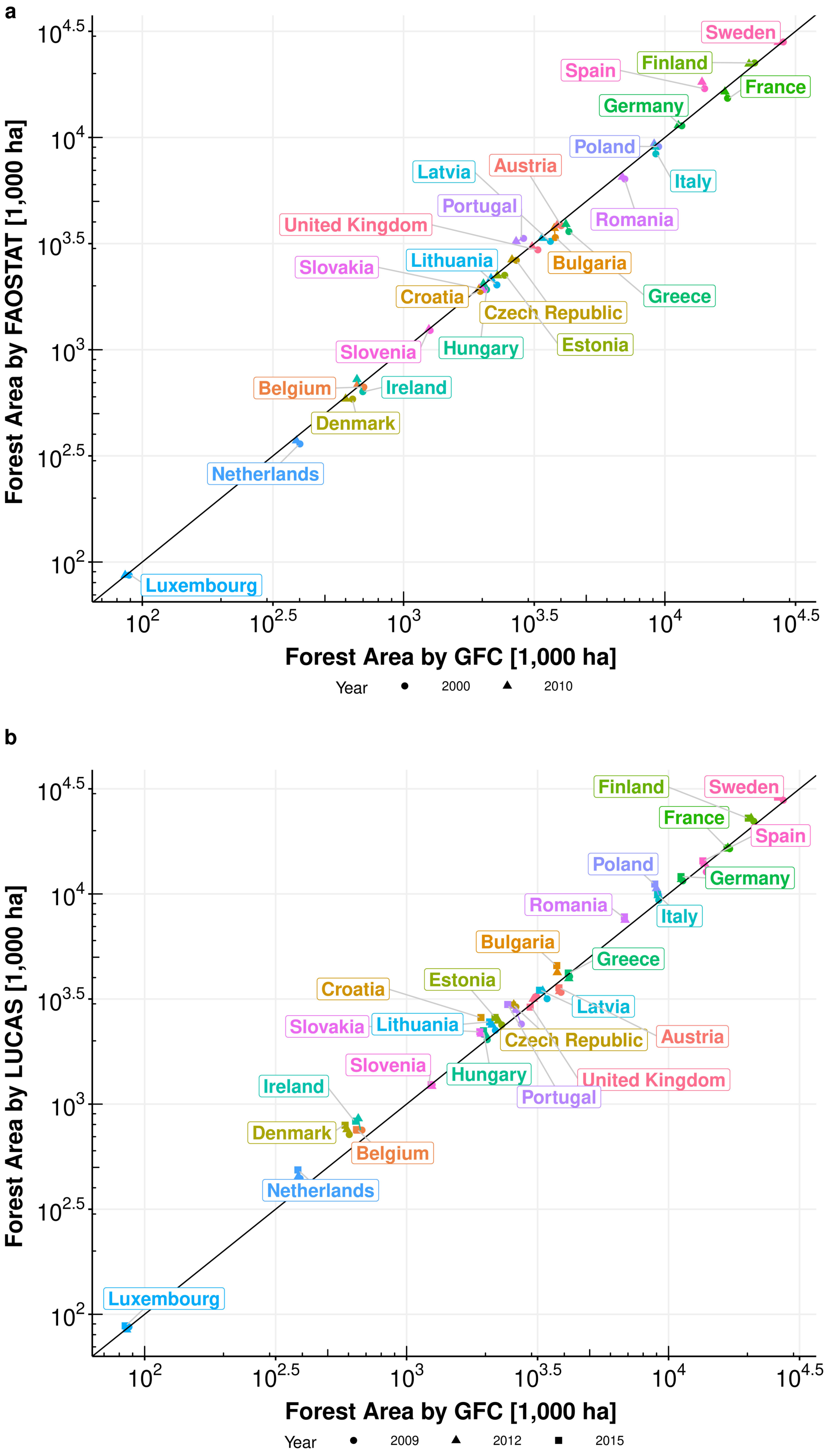
Author contributions G.C. and A.C. conceived the idea and designed the methodology; G.C. analysed the data and wrote the Google Earth Engine and R scripts; G.C. and A.C. wrote the manuscript with contributions from G.D., G.L., V.A., R.P. and G.G.. All authors contributed critically to the interpretation of the results and gave final approval for publication.

Competing interests The authors declare no competing interests.

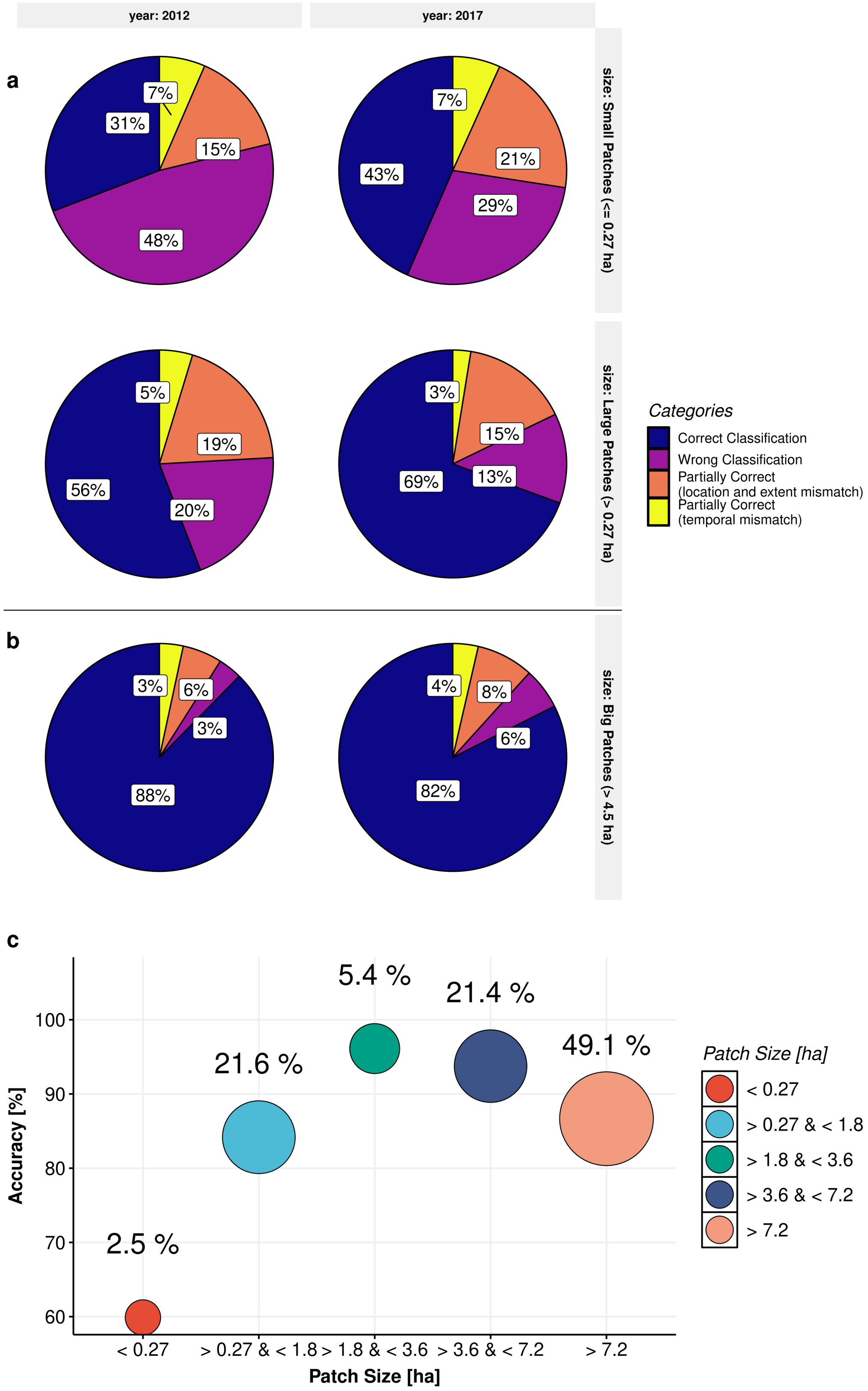
Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2438-y.
Correspondence and requests for materials should be addressed to G.C.
Peer review information *Nature* thanks Mathew Williams and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



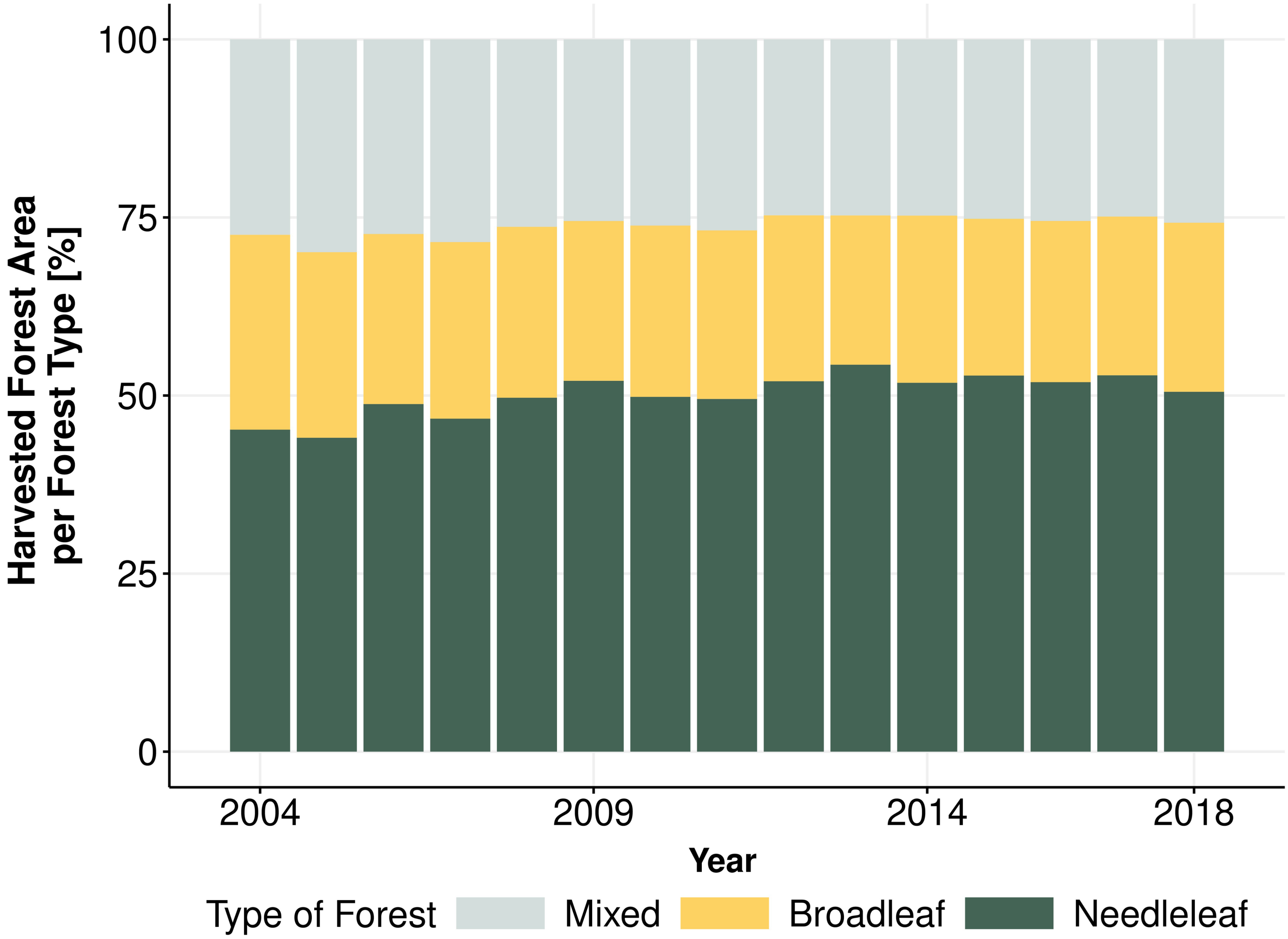
Extended Data Fig. 1 | From tree cover to forest cover. a, Tree-cover threshold needed to define a forest (colours) and percentage error between FAOSTAT-2015 and remote-sensing-based forests (labels). **b**, Forest threshold sensitivity. Maps were generated using GEE[22].



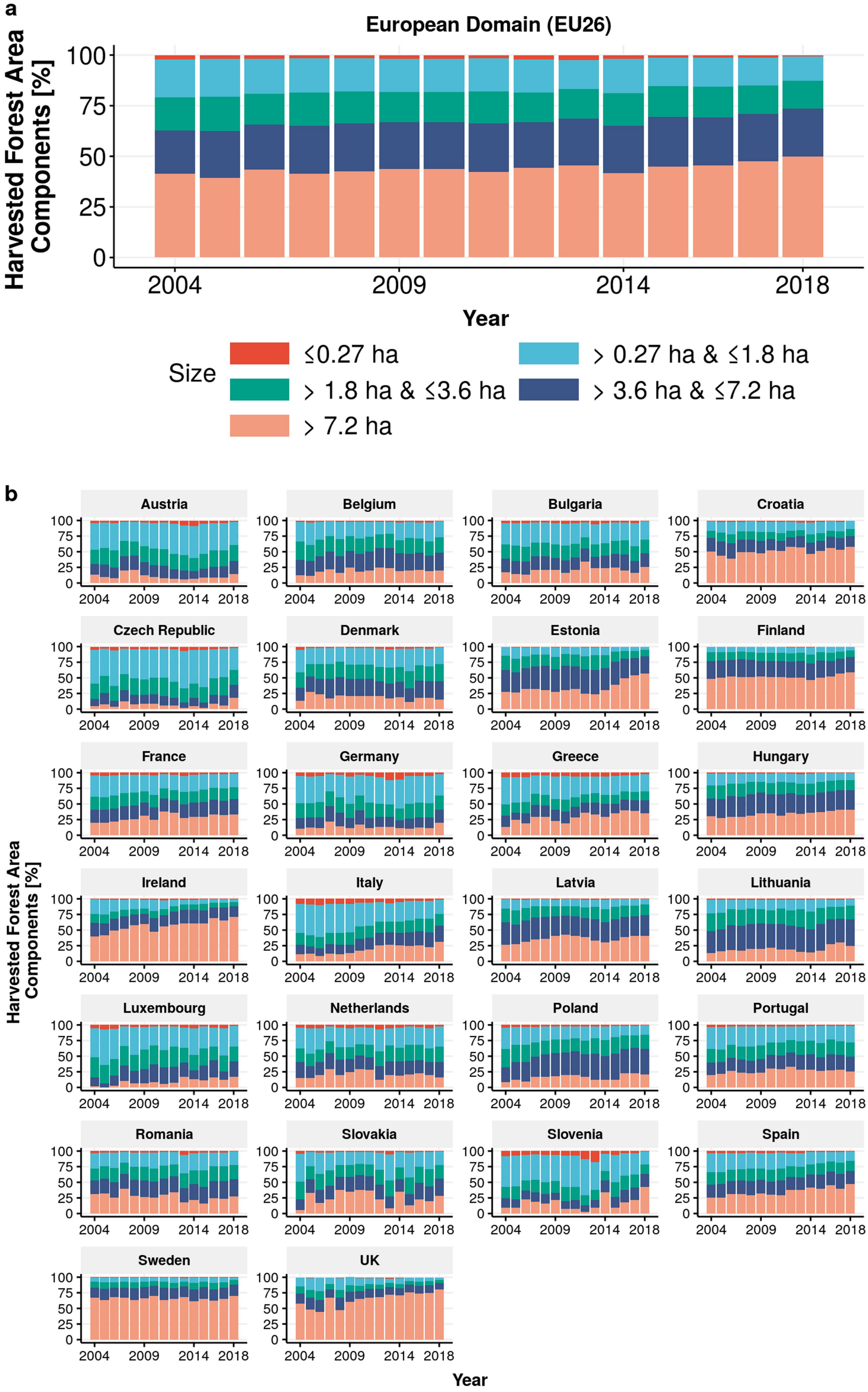
Extended Data Fig. 2 | Verification of EU forest area. a, GFC data versus FAOSTAT for 2000 and 2010. **b**, GFC data versus LUCAS for 2009, 2012 and 2015.



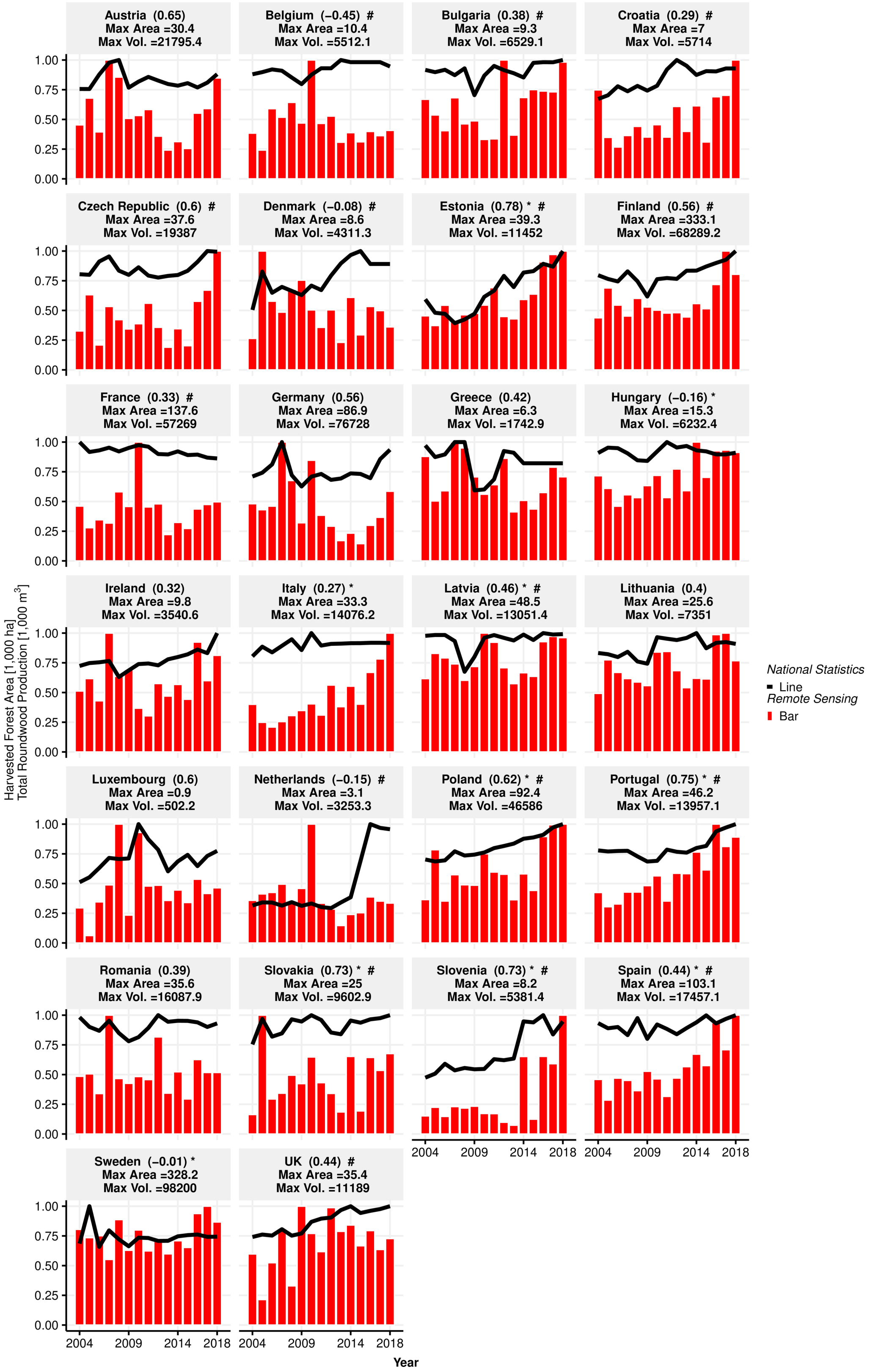
Extended Data Fig. 3 | Validation of GFC-derived forest loss with high-resolution data. a, b, Validation of the classification of harvested areas in the years 2012 and 2017 by forest patches of sizes small (≤0.27 ha) and large (>0.27 ha and ≤4.5 ha; **a**), and big (>4.5 ha; **b**). **c**, Accuracy of harvest area derived from GFC-derived forest loss versus patch size (labels and circle size refer to the EU26-wise cumulative harvested forest).



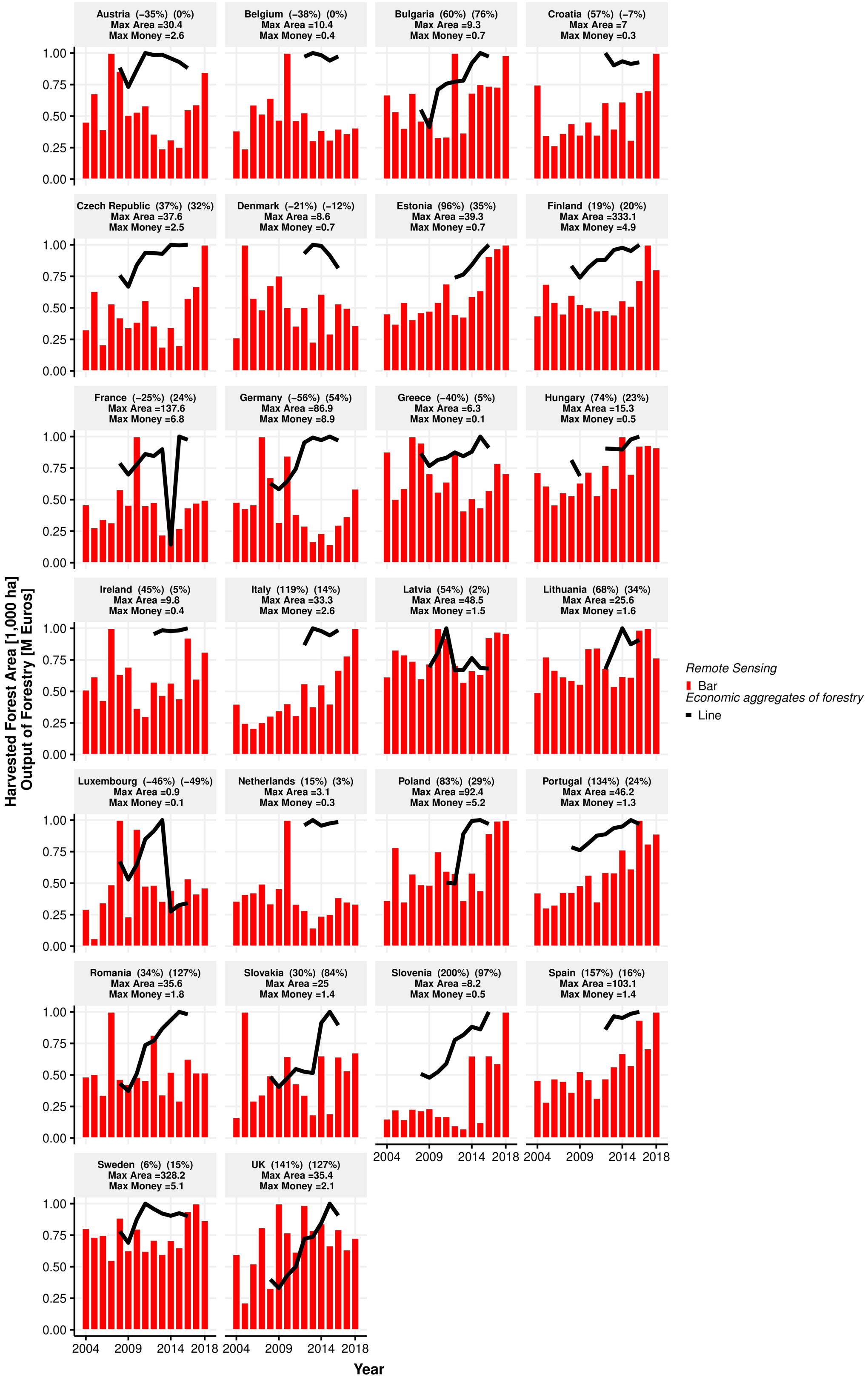
Extended Data Fig. 4 | Harvested forest area by forest type. Time series of land cover type (from GlobCover)[25] for EU26. Colours refer to the three forest types: mixed, broadleaf and needleleaf.



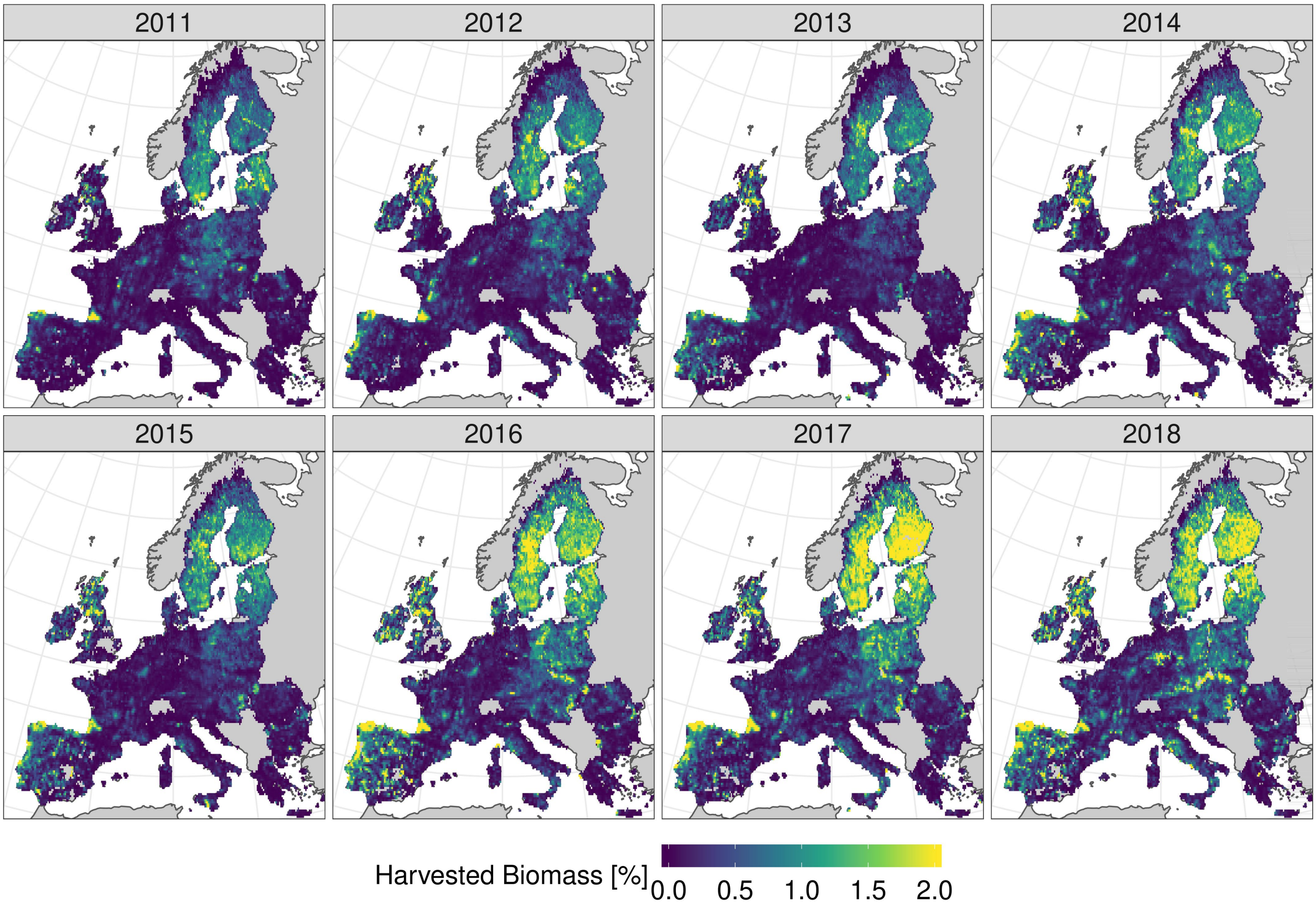
Extended Data Fig. 5 | Harvested forest area components. a, **b**, Annual distribution of harvested forest for different classes of patch size, ranging from small patches (harvested forest area less than 0.27 ha) to big patches (harvested forest area greater than 7.2 ha) for all of EU26 (**a**), each EU26 country (**b**).



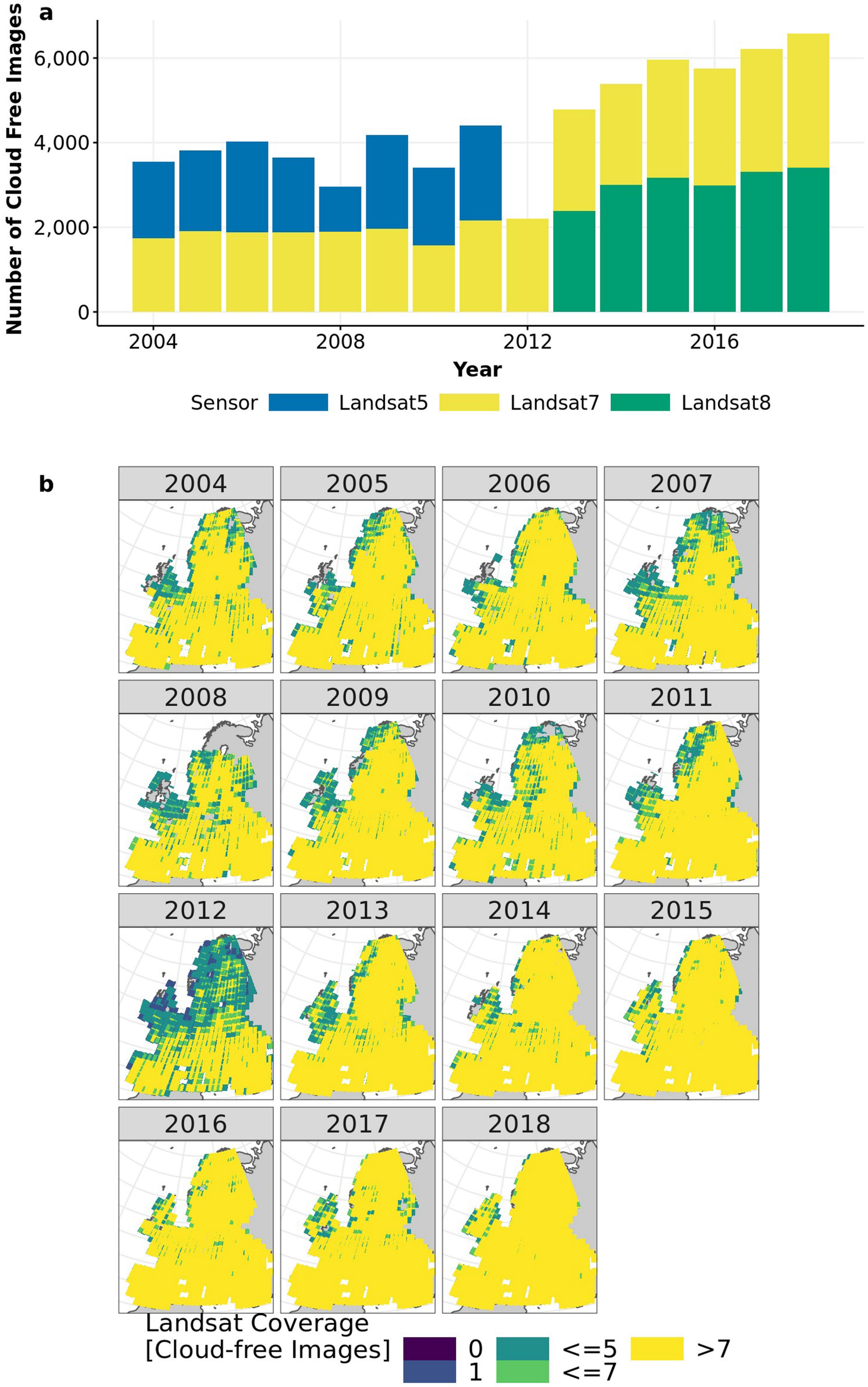
Extended Data Fig. 6 | GFC-derived harvested forest area versus official harvest removal data. Harvested forest area from the GFC maps (red bars, normalized between 0 and 1) and volumes of harvest removals from national statistics (black lines, normalized between 0 and 1). We excluded areas affected by forest fires and retained areas affectedby major windstorms because they appear in the harvest removal data. Statistical significance at $P = 0.05$ for remote sensing and national statistics is indicated by an asterisk and a hash, respectively, in the country label panels. The value in brackets is the correlation coefficient, r. Maximum values of harvested forest area and official harvest removal data for each country are reported in the second and third lines of each label, respectively.



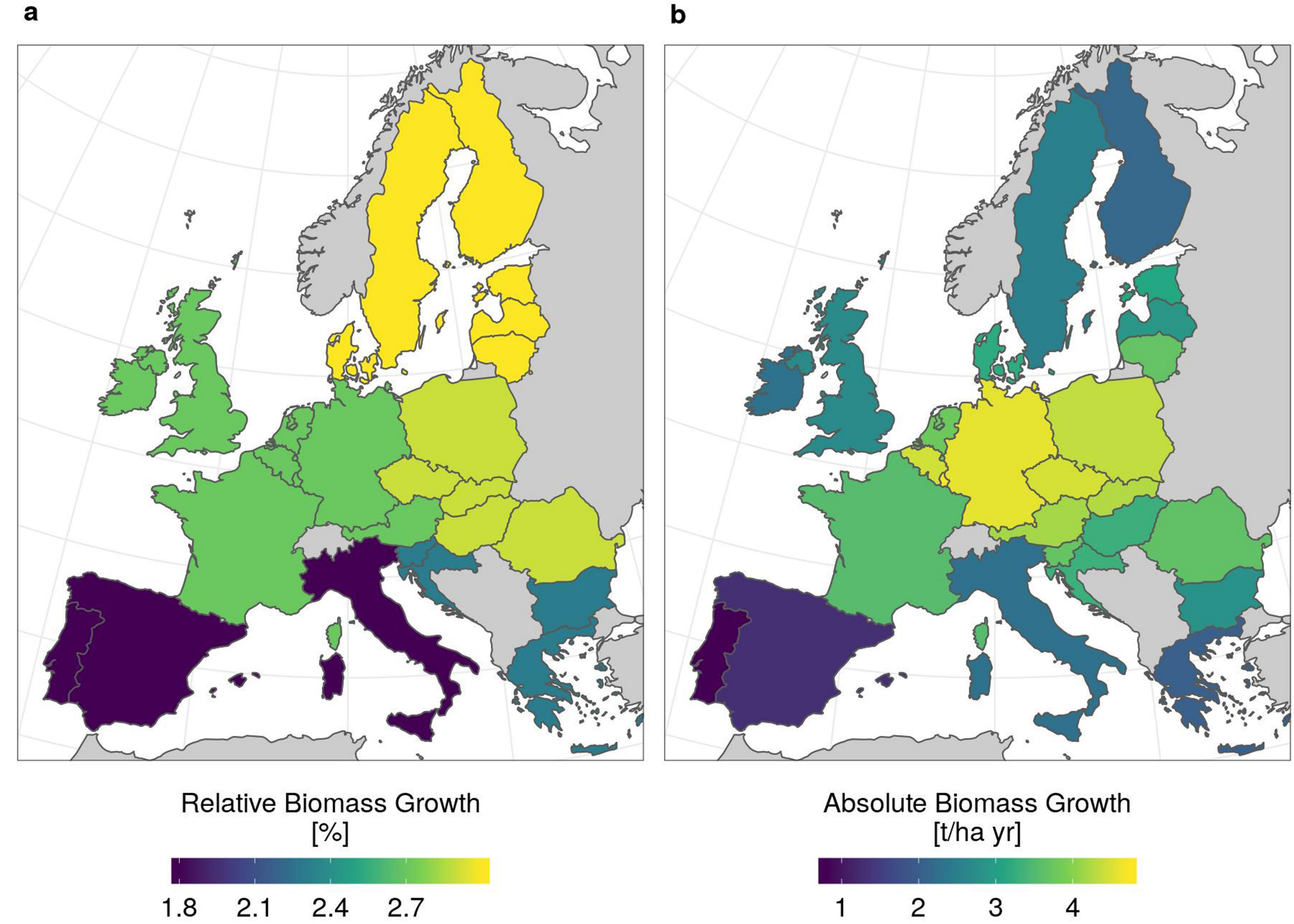
Extended Data Fig. 7 | Harvested forest area versus Eurostat[32] economic aggregates. Harvested forest area from the GFC maps (red bars, normalized between 0 and 1) and volumes of economic aggregates of forestry from Eurostat data (black lines, normalized between 0 and 1). We excluded areas affected by forest fires and retained areas affected by major windstorms because they appear in the harvest removal data. Percentages in the first and second brackets after the country label refer to the percentage change 2008–2016 (or 2012–2016 when 2008 records are not available) of remote sensing and market value, respectively. Maximum values of harvested forest area and volumes of economic aggregates of forestry for each country are reported in the second and third lines of each label, respectively.



Extended Data Fig. 8 | Harvested forest biomass per year. Percentage of AGB harvested (expressed as relative amount of biomass affected by management practices) per year in a 0.2° grid cell excluding forest losses due to fires and major windstorms and areas with sparse forest cover. As in Fig. 1 but measuring biomass instead. This map was generated using GEE[22].



Extended Data Fig. 9 | Cloud-free land coverage of Landsat in Europe. **a**, Time series of cloud-free Landsat scenes (cloud cover less than 20%) for EU26. **b**, Spatial distribution of cloud-free Landsat images over Europe. Grey areas indicate where no data was available for the selected year using satellite imagery. Map and time series were generated using GEE[22].



Extended Data Fig. 10 | Growth rates of forest biomass. a, **b**, Relative (**a**) and absolute (**b**) growth rate of forest biomass as derived from the *State of Europe's Forests 2015* report[1] in combination with GlobBiomass[27] and GFC[21] data. The data in **a** are given over five European regions, with colours corresponding to the colour scale: north (yellow), central west (green), central east (lime), south west (purple) and south east (blue). Maps were generated using GEE[22].

Revealing enigmatic mucus structures in the deep sea using DeepPIV

https://doi.org/10.1038/s41586-020-2345-2

Received: 26 June 2019

Accepted: 19 March 2020

Kakani Katija[1]✉, Giancarlo Troni[2], Joost Daniels[1], Kelly Lance[1], Rob E. Sherlock[1], Alana D. Sherman[1] & Bruce H. Robison[1]

Many animals build complex structures to aid in their survival, but very few are built exclusively from materials that animals create [1,2]. In the midwaters of the ocean, mucoid structures are readily secreted by numerous animals, and serve many vital functions[3,4]. However, little is known about these mucoid structures owing to the challenges of observing them in the deep sea. Among these mucoid forms, the 'houses' of larvaceans are marvels of nature[5], and in the ocean twilight zone giant larvaceans secrete and build mucus filtering structures that can reach diameters of more than 1 m[6]. Here we describe in situ laser-imaging technology[7] that reconstructs three-dimensional models of mucus forms. The models provide high-resolution views of giant larvacean houses and elucidate the role that house structure has in food capture and predator avoidance. Now that tools exist to study mucus structures found throughout the ocean, we can shed light on some of nature's most complex forms.

In the growing field of bioinspired design, many technological innovations have been developed from forms found in nature. Whereas some animals build structures to protect themselves from the elements and predation, other animals build structures that enable important functions such as feeding[1]. Although many of these constructs are fabricated from found materials, a few animals (for example, terrestrial spiders and some marine fish and invertebrates) secrete specialized compounds that they fashion into complex structures[2,3,5]. In the midwaters of the ocean, mucoid structures are readily secreted by a variety of animals, and can serve many vital functions including food collection[4] and protection from pathogens and predators[3,6,8]. Larvaceans (also known as appendicularians) are pelagic tunicates that secrete and inhabit mucus structures (or 'houses'), in which rhythmic pumping of the tail of the larvacean drives water-borne particles through filters to concentrate food into the mouth of the larvacean[5,9–12]. As the second-most abundant mesozooplankton in the ocean[13–16], larvaceans with their mucus houses are thus able to feed on particles and prey from sub-micrometre to sub-millimetre size scales (for example, bacteria to microzooplankton), thereby broadly influencing oceanic food webs[15]. The remarkable complexity of these houses—and their effective mechanisms for the selection and retention of particles[17,18]—could provide promising leads for future bioinspired designs of pumping and filtration systems.

Larvacean sizes range by an order of magnitude. Among the three families within the class Appendicularia (the Fritillariidae, Kowalevskiidae and Oikopleuridae), species differ in body size, morphology and house structure[12]. Smaller oikopleurids are typically less than 1 cm in body length and their houses are approximately two times larger[12]. Larger oikopleurids—known as giant larvaceans (which include the genus *Bathochordaeus*)—can be up to 10 cm in length[19], and their mucus houses can be up to a metre in the largest dimension[6]. Giant larvaceans occur globally[20], and three different species of *Bathochordaeus* (*Bathochordaeus charon*, *Bathochordaeus stygius* and *Bathochordaeus mcnutti*)[19,21] occur in Monterey Bay, where the more abundant *B. mcnutti* and *B. stygius* occur stratified vertically in the water column[19,22]. Giant larvaceans are important players in the biological pump of the ocean[23] and have previously been shown to directly contribute to carbon cycling in the deep sea[7,22–24].

Even though giant larvaceans are ecologically important, little is known of the internal structure and flow pathways within their houses because of the challenges of conducting long-duration observations in the midwaters of the ocean, the lack of quantitative in situ observational tools and the difficulty in maintaining captive animals. Sampling with nets typically results in the collection of larvaceans without their delicate filtration structures, and the association of giant larvaceans with their mucus houses was not established until in situ observations were made[25]. Early investigations of house building and filter feeding by shallow, smaller larvaceans such as the genus *Oikopleura* were based on hand- or net-collected specimens observed in the laboratory[10,26] and later kept in culture[27–30]. With the advent of manned and robotic exploration, the ecology of giant larvacean houses has been described in further detail[6,23,25,31]. However, investigations continue to be limited by the observational tools available, and attempts to observe the generation and structure of houses of large, deep-living larvaceans in the laboratory have been unsuccessful.

To understand the function of giant larvacean houses, their structure must be resolved. Houses consist of a coarse-mesh outer house, which surrounds a complex, fine-mesh inner house that contains a food-concentrating filter (Fig. 1a). The filtering structures of all larvaceans are constructed of transparent mucopolysaccharides[32,33] including specific oikosin structural proteins that—in *Oikopleura dioica*—interact with a scaffold of cellulose microfibrils[27,34]. Apart from size, there are appreciable differences between the structure and

---

[1]Research and Development, Monterey Bay Aquarium Research Institute, Moss Landing, CA, USA. [2]Department of Mechanical Engineering, Pontificia Universidad Católica de Chile, Santiago, Chile. ✉e-mail: kakani@mbari.org

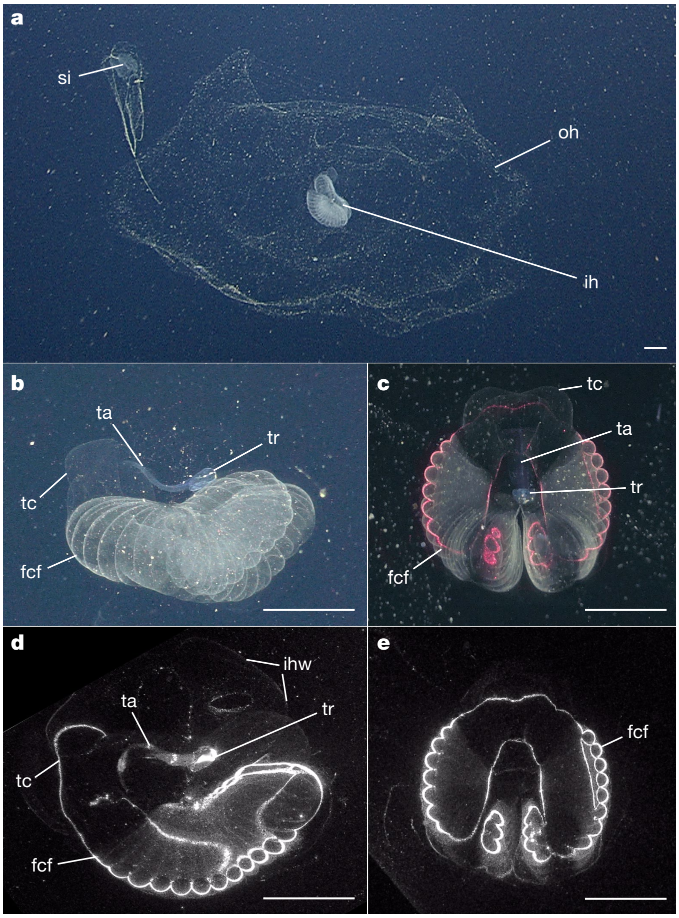


Fig. 1 | Giant larvacean, *B. stygius*, in its mucus feeding structure, which includes an inner and outer house. a, Inner and outer house structures of the mucus feeding structure. **b**–**e**, White-light (**b**, **c**) and laser-sheet (**d**, **e**) illumination of both the lateral (**b**, **d**) and dorsal (**c**, **e**) views of a midwater giant larvacean, *B. stygius*. fcf, food-concentrating filter; ih, inner house; ihw, internal house wall; oh, outer house; si, abandoned house or sinker; st, suspensory thread; ta, tail; tc, tail chamber; tr, trunk. Scale bars, 4 cm.

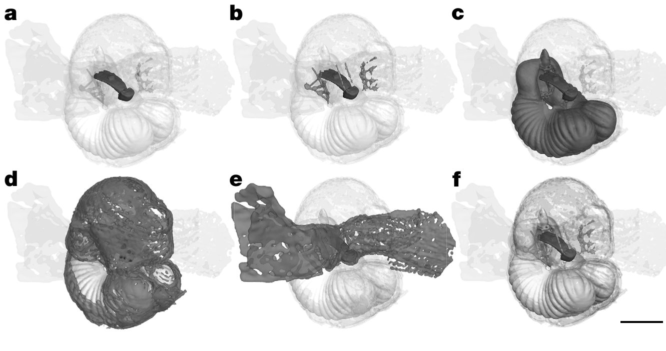


Fig. 2 | A three-dimensional reconstructed model of a giant larvacean and its inner house yields composite models of the mucus structure. a–**e**, The reconstructed model enables the visualization of distinct features (dark grey) that include the animal (**a**; black in all panels), the inlet filters and suspensory threads adjacent to the animal (**b**), the food-concentrating filter and buccal tube (**c**), the internal house wall (**d**) and the inlet channels to the inner house (**e**). **f**, The entire inner house model. Scale bar, 4 cm.

function of the houses of giant larvaceans and smaller oikopleurids[12], and two reports on *Bathochordaeus* houses have differing views on the structure of the inner house and production of the outer house[6,12]. In situ feeding experiments using remotely operated vehicles (ROVs) have demonstrated that giant larvaceans can ingest particles ranging in size from 10 to 600 µm in diameter[22]. Further observations with new in situ imaging tools are needed to resolve the differences in house morphology, and to elucidate the means by which giant larvaceans select and process particles.

Here we present three-dimensional in situ visualizations of the mucus structures of giant larvaceans using an ROV-deployable laser-sheet imaging device called DeepPIV. While operating in stationary mode, DeepPIV is used to study small-scale, particle–fluid interactions[7,22]. Operated in scanning mode, the laser sheet of DeepPIV penetrates cleanly through gelatinous and mucus structures, and cross-sectional images of these structures can be collected through the translational fore or aft motion of the ROV. The resulting image stacks are then used to generate three-dimensional reconstructions that provide new views of gelatinous or mucus structures, and in this case *Bathochordaeus* within its house. Together with ever-advancing microscopy and molecular methods[28,30,35–37], this in situ observational tool elucidates meso- and microscale structures and fluid flow.

Giant larvaceans were observed from June to December 2015 during cruises on RVs *Western Flyer* and *Rachel Carson* in Monterey Bay, California, USA. DeepPIV was deployed on 13 separate dives affixed to the Monterey Bay Aquarium Research Institute's ROV *MiniROV* (Extended Data Fig. 1), a fly-away vehicle that is rated to 1,500 m. During deployments between 100- and 400-m depths, 71 specimens of the genus *Bathochordaeus* were observed (Fig. 1). Of these, three-dimensional reconstruction measurements were conducted on 14 individual *B. stygius* houses (Supplementary Videos 1, 2). After using strict criteria for the selection of three-dimensional reconstruction data (Supplementary Information), 6 in situ reconstruction datasets, corresponding to 5 individuals that ranged in trunk length from 0.4 to 2.1 cm, were used for subsequent analysis (Supplementary Table 1). Here we use descriptors (for example, dorsal and ventral) to refer to the most common orientation of the mucus house when found in midwater.

Models reveal structures in giant larvacean houses

Using the DeepPIV-derived image scans of occupied *B. stygius* inner houses, the three-dimensional structure was determined (Supplementary Information), yielding composite views of house features (Figs. 2, 3, Extended Data Fig. 2 and Supplementary Video 3). Isolation of features by pixel intensity and other metrics enabled the identification of specific structures, including some that have not previously been described or that are contested in the literature[6,12] (Fig. 2). The three-dimensional reconstruction models also provide interior views of the inner house (Figs. 3, 4 and Extended Data Fig. 2). These newly resolved components include upper cushion chambers, ramp valves, escape chambers, and the interface between the food-concentrating filters and supply chambers (Figs. 3, 4). Given the lack of structure at the exit of the exhaust passage (Fig. 3e), we suspect that an exhaust valve is not present and that, therefore, a giant larvacean cannot propel its entire house, as some smaller larvaceans do[12,38,39]. Finally, a particle-rejection passageway from the buccal tube to an exhaust chamber—similar to that observed in *O. dioica*[35]—could not be identified. Owing to its exceedingly fragile nature, we were unable to conduct full three-dimensional reconstruction scans of the outer house, and a conceptual drawing based on high-resolution videos of the outer and inner house is shown in Extended Data Fig. 3.

We coupled three-dimensional reconstructions with dye visualizations (Supplementary Video 4) and measurements of particle movement using DeepPIV (Supplementary Video 5) to gain insights into the interior flow patterns and how the mucus structure functions as a food-concentrating filter (Fig. 4). Water flow enters the inner house through two inlet channels that connect to the outer house. At the entrance to the inner house are inlet filters, with an average rectangular

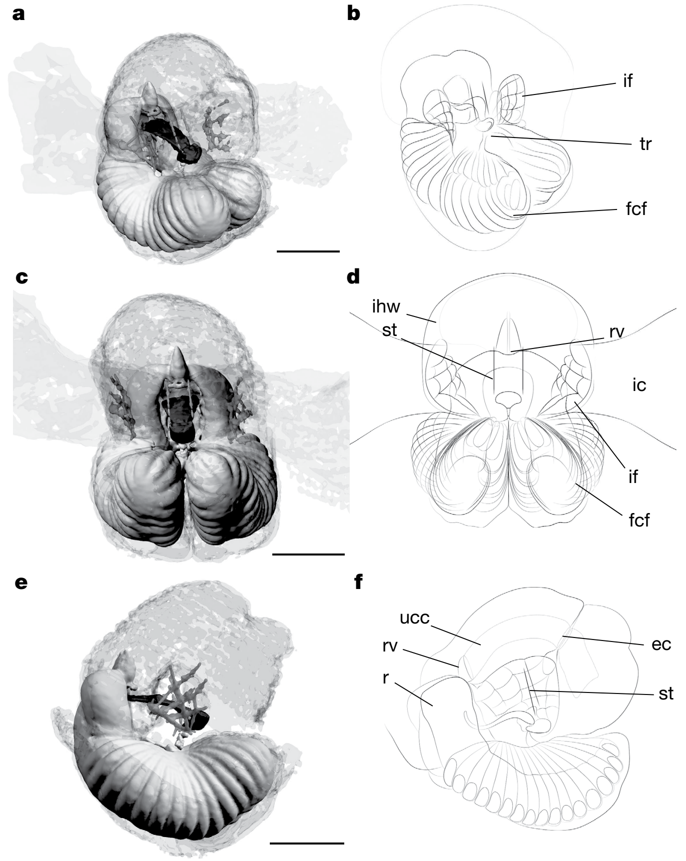


Fig. 3 | Comparison between traditional line sketches and three-dimensional models of a giant larvacean mucus house. **a**–**f**, Isometric (**a**, **b**), anterior (**c**, **d**) and lateral (**e**, **f**) views of three-dimensional reconstructed models (**a**, **c**, **e**) and traditional line sketches (**b**, **d**, **f**) of a giant larvacean inside its mucus house. The larvacean is shown in black in the three-dimensional model. The inlet channel is not shown in **b**, **e**, **f**; the outer house is not shown in any of the images. ec, escape chamber; ic, inlet channel; if, inlet filter; r, ramp; rv, ramp valve; ucc, upper cushion chamber. Scale bars, 4 cm.

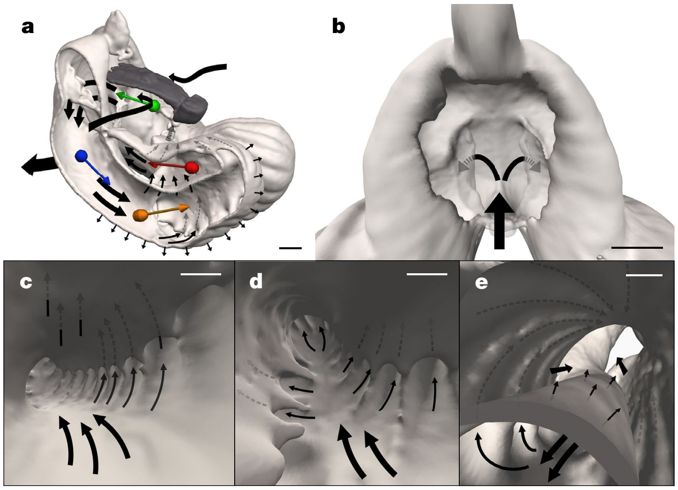


Fig. 4 | Laser scanning coupled with particle flow-field measurements reveal the structure and function of the mucus house. **a**, Generalized flow direction (indicated by black arrows weighted by flow volume; dashed grey arrows show flow in occluded passageways) inside the larvacean mucus house. **b**–**e**, Views at different points within the mucus house—as indicated by coloured dots and arrows in **a**—show finer details from within the tail chamber (**b**; green dot and arrow in **a**), the posterior end of the main inner filter chamber (**c**; blue dot and arrow in **a**), the anterior end of the main inner filter chamber (**d**; orange dot and arrow in **a**) and between the main inner filter chamber and exit chamber (**e**; red dot and arrow in **a**). Scale bars, 1 cm.

mesh size of 3.6 ± 1.7 mm (Supplementary Table 1). After the inlet filters, flow enters the tail chamber where the larvacean resides. The giant larvacean body is surrounded by mucus, with closely fitting tail chamber walls, which enable effective momentum transfer between the pumping larvacean tail and the fluid[12]. Flow moves through the tail chamber and bifurcates into two separate supply passages, leading to the food-concentrating filter. We observed flow passing through the inner and outer surfaces of the ridges of the food-concentrating filter, as well as through the ceilings of the supply chambers, which then lead to a central exhaust channel that redirects flow towards the posterior end of the house (Fig. 4c–e and Supplementary Videos 4, 5) and empties into the interior of the outer house. Once water is excluded from these regions, dense particle suspensions continue to move through the ridges of the food-concentrating filter toward the buccal tube and into the mouth of the larvacean (see also supplementary video 1 of a previous study[22]).

Although only partially resolved in the three-dimensional reconstruction model (Fig. 3 and Supplementary Video 3), we found additional features that provide clues about the specific functions of internal features (Supplementary Videos 4, 5). A ramp valve (Fig. 3d, f) at the top of the inner house regulates flow through the upper cushion chamber (Fig. 3f) and connects with the tail and body chambers (Fig. 3f). In addition, paired valves at the exit of the tail chamber (Fig. 4b) may regulate the flow into the supply passages and food-concentrating filter. Along with the suspensory threads (Figs. 2b, 3b, d, f), these chambers and valves within the inner house wall probably have a role in the maintenance of structural rigidity; the connection between these chambers and the tail chamber enables the maintenance of hydrostatic pressure. An escape chamber (Fig. 3f), which cuts through the upper cushion chambers, is located above the trunk of the animal. Additional lateral chambers were also observed but are not shown, owing to the lack of detail.

Ecological functions of the mucus structures

The visualizations of mucus house structures have highlighted key differences between the houses of giant larvaceans and smaller oikopleurids that are probably rooted in differences in ecological function. Although the mesh 'external walls' found in *Oikopleura labradoriensis*[10] could be analogous to the outer house of *B. stygius*, the sheer size of the outer house relative to the inner house structure, as well as the presence of inlet channels, are features that are not found in *O. labradoriensis* and may be unique to *Bathochordaeus*. The outer house probably provides a coarse-mesh shield that prevents large particles from compromising the fine-mesh inner house. Estimates show that giant larvaceans regenerate a house approximately every day[24]—and together with the physiological status of the animal[40]—this replacement rate may be due in part to damage inflicted by diel vertical migrating animals.

The mucus house of giant larvaceans could also serve as a physical barrier to deter predation. The internal house wall and outer house may prevent larvacean contact with stinging cells of abundant gelatinous predators such as medusae and siphonophores[6]. Additionally, fishes are known to prey on larvaceans[41], using sensory organs[42] that allow the fish to detect hydromechanical cues that correspond to escape jumps of copepods and other small mesozooplankton prey[43,44]. For such mechanosensing predators, the outer house may act as a cloaking device to hydrodynamically shield the filter-feeding giant larvacean. To quantify this effect, we can model the outer house as a mesh sphere with radius r, in which the surface area $A_h = 4\pi r^2$ is related to the inlet and outlet surface area of the outer house (A_i and A_o, respectively) by $A_h = A_i + A_o$. By assuming that flow is steady in the mucus house, and the outer house radius is constant once fully expanded by the larvacean, the relationship between the flow generated by the animal through the tail chamber (u_t), the flow entering the outer house (u_i) and the flow escaping through the outer house (u_o) can be given by the continuity equation[45]

$$u_t A_t = u_o A_o = u_i A_i,$$

in which A_t is the cross-sectional area of the tail chamber. Solving for u_o gives

$$u_o = \frac{u_t A_t}{4\pi r^2 \eta},$$

in which $\eta = \frac{A_o}{A_h}$, and the precise value of A_o is unknown. On the basis of our observations (Extended Data Fig. 3), A_o may vary from 50 to 90% of A_h, and most probably lies at the upper end of that range. Using average values for u_t and A_t of 1.33 cm s^{-1} and 4.82 cm^2, respectively, for *B. stygius*[7], and an *r* of 50 cm for the outer house[24], the flow escaping the outer house u_o is estimated to be 3,300 to 5,800 times smaller than u_t. Therefore, instead of detecting the flow induced by the beating tail of a giant larvacean, a passing predator will instead encounter flow escaping the mucus house that is three orders of magnitude smaller. The resulting reduction in flow could lower the probability of detection by mechanosensing predators, making predation on filter-feeding giant larvaceans less likely.

Complexity of mucus houses as a marvel of nature

The mucus feeding houses of larvaceans, both small and large, are marvels of intricate complexity. First described around the turn of the twentieth century, house morphology and function were largely unknown until Alldredge used SCUBA to describe the houses of seven species of small oikopleurids[32]. The in situ observations by Alldredge revealed pathways of water flow through filters that maximize the concentration and retention of particles. Similarly, the basic inner and outer house structure of the deep-living giant larvaceans was first resolved by in situ observations from an ROV[6]. The detailed internal structure and water flow patterns through the chambers, filters and valves of the houses of giant larvaceans have now been revealed. The research that enabled this advance was conducted in situ using DeepPIV, which will make the study of other gelatinous and mucus structures in midwater possible (Extended Data Fig. 4).

The greatest remaining mysteries of larvacean houses concern how they are produced. Whereas a spider builds a complicated web one silky strand at a time[1,2], the house of a larvacean is extruded all at once as a rudiment and is then inflated[5]. This leads to the question of how a bank of mucus-producing cells can create such an intricate form within a small, tightly packed bubble. Given their remarkable architecture, it seems almost implausible that these complex marvels should be built to last only a day or two[24]. Future observational tools and vehicles will enable us to observe the construction of giant larvacean houses in their entirety, and to precisely document the frequency with which they are built.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2345-2.

1. Hansell, M. *Built by Animals: The Natural History of Animal Architecture* (Oxford Univ. Press, 2007).
2. Gosline, J. M., DeMont, M. E. & Denny, M. W. The structure and properties of spider silk. *Endeavour* **10**, 37–43 (1986).
3. Grutter, A. S., Rumney, J. G., Sinclair-Taylor, T., Waldie, P. & Franklin, C. E. Fish mucous cocoons: the 'mosquito nets' of the sea. *Biol. Lett.* **7**, 292–294 (2011).
4. Gilmer, R. W. Free-floating mucus webs: a novel feeding adaptation for the open ocean. *Science* **176**, 1239–1240 (1972).
5. Alldredge, A. L. Appendicularians. *Sci. Am.* **235**, 94–105 (1976).
6. Hamner, W. M. & Robison, B. H. In situ observations of giant appendicularians in Monterey Bay. *Deep Sea Res. A* **39**, 1299–1313 (1992).
7. Katija, K., Sherlock, R. E., Sherman, A. D. & Robison, B. H. New technology reveals the role of giant larvaceans in oceanic carbon cycling. *Sci. Adv.* **3**, e1602374 (2017).
8. Ellis, A. E. Innate host defense mechanisms of fish against viruses and bacteria. *Dev. Comp. Immunol.* **25**, 827–839 (2001).
9. Fol, H. Etudes sur les Appendiculaires du détroit de Messine. *Mem. Soc. Phys. Hist. Nat. Geneve* **21**, 445–499 (1872).
10. Flood, P. R. Architecture of, and water circulation and flow rate in, the house of the planktonic tunicate *Oikopleura labradoriensis*. *Mar. Biol.* **111**, 95–111 (1991).
11. Acuña, J. L., Deibel, D. & Morris, C. C. Particle capture mechanism of the pelagic tunicate *Oikopleura vanhoefeni*. *Limnol. Oceanogr.* **41**, 1800–1814 (1996).
12. Flood, P. R. & Deibel, D. in *The Biology of Pelagic Tunicates* 105–125 (Oxford Univ. Press, 1998).
13. Landry, M. R., Peterson, W. K. & Fagerness, V. L. Mesozooplankton grazing in the Southern California Bight. I. Population abundances and gut pigment contents. *Mar. Ecol. Prog. Ser.* **115**, 55–71 (1994).
14. Hopcroft, R. R. & Roff, J. C. Production of tropical larvaceans in Kingston Harbour, Jamaica: are we ignoring an important secondary producer? *J. Plankton Res.* **20**, 557–569 (1998).
15. Gorsky, G. & Fenaux, R. in *The Biology of Pelagic Tunicates* 161–169 (Oxford Univ. Press, 1998).
16. Jaspers, C., Nielsen, T. G., Carstensen, J., Hopcroft, R. R. & Moller, E. F. Metazooplankton distribution across the Southern Indian Ocean with emphasis on the role of larvaceans. *J. Plankton Res.* **31**, 525–540 (2009).
17. Fernández, D., Lopez-Urrutia, A., Fernández, A., Acuña, J. L. & Harris, R. P. Retention efficiency of 0.2 to 6 μm particles by the appendicularians *Oikopleura dioica* and *Fritillaria borealis*. *Mar. Ecol. Prog. Ser.* **266**, 89–101 (2004).
18. Conley, K. R., Lombard, F. & Sutherland, K. R. Mammoth grazers on the ocean's minuteness: a review of selective feeding using mucous meshes. *Proc. R. Soc. B* **285**, 20180056 (2018).
19. Sherlock, R. E., Walz, K. R., Schlining, K. L. & Robison, B. H. Morphology, ecology, and molecular biology of a new species of giant larvacean in the eastern North Pacific: *Bathochordaeus mcnutti* sp. nov. *Mar. Biol.* **164**, 20 (2017).
20. Flood, P. R. in *Response of Marine Ecosystems to Global Change: Impact of Appendicularians* (eds Gorsky, G. et al.) 59–85 (Contemporary Publishing International, 2005).
21. Sherlock, R. E., Walz, K. R. & Robison, B. H. The first definitive record of the giant larvacean, *Bathochordaeus charon*, since its original description in 1900 and a range extension to the northeast Pacific Ocean. *Mar. Biodivers. Rec.* **9**, 79 (2016).
22. Katija, K., Choy, C. A., Sherlock, R. E., Sherman, A. D. & Robison, B. H. From the surface to the seafloor: how giant larvaceans transport microplastics into the deep sea. *Sci. Adv.* **3**, e1700715 (2017).
23. Silver, M. W., Coale, S. L., Pilskaln, C. H. & Steinberg, D. R. Giant aggregates: importance as microbial centers and agents of material flux in the mesopelagic zone. *Limnol. Oceanogr.* **43**, 498–507 (1998).
24. Robison, B. H., Reisenbichler, K. R. & Sherlock, R. E. Giant larvacean houses: rapid carbon transport to the deep sea floor. *Science* **308**, 1609–1611 (2005).
25. Barham, E. G. Giant larvacean houses: observations from deep submersibles. *Science* **205**, 1129–1131 (1979).
26. Deibel, D. Feeding mechanism and house of the appendicularian *Oikopleura vanhoeffeni*. *Mar. Biol.* **93**, 429–436 (1986).
27. Sagane, Y., Hosp, J., Zech, K. & Thompson, E. M. Cytoskeleton-mediated templating of complex cellulose-scaffolded extracellular structure and its association with oikosins in the urochordate *Oikopleura*. *Cell. Mol. Life Sci.* **68**, 1611–1622 (2011).
28. Hosp, J., Sagane, Y., Danks, G. & Thompson, E. M. The evolving proteome of a complex extracellular matrix, the *Oikopleura* house. *PLoS ONE* **7**, e40172 (2012).
29. Vaugeois, M., Diaz, F. & Carlotti, F. A mechanistic individual-based model of the feeding processes for *Oikopleura dioica*. *PLoS ONE* **8**, e78255 (2013).
30. Martí-Solans, J. et al. *Oikopleura dioica* culturing made easy: a low-cost facility for an emerging animal model in EvoDevo. *Genesis* **53**, 183–193 (2015).
31. Hopcroft, R. R. & Robison, B. H. A new mesopelagic larvacean, *Mesochordaeus erythrocephalus*, sp. nov., from Monterey Bay, with a description of its filtering house. *J. Plankton Res.* **21**, 1923–1937 (1999).
32. Alldredge, A. L. House morphology and mechanisms of feeding in the Oikopleuridae (Tunicata, Appendicularia). *J. Zool.* **181**, 175–188 (1977).
33. Körner, W. F. Untersuchungen über die gehäusebildung bei appendicularien (*Oikopleura dioica* Fol). *Z. Morphol. Oekol. Tiere* **41**, 1–53 (1952).
34. Kishi, K., Hayashi, M., Onuma, T. A. & Nishida, H. Patterning and morphogenesis of the intricate but stereotyped oikoplastic epidermis of the appendicularian, *Oikopleura dioica*. *Dev. Biol.* **428**, 245–257 (2017).
35. Conley, K. R., Gemmell, B. J., Bouquet, J.-M., Thompson, E. M. & Sutherland, K. R. A self-cleaning biological filter: how appendicularians mechanically control particle adhesion and removal. *Limnol. Oceanogr.* **63**, 927–938 (2018).
36. Flood, P. R., Deibel, D. & Morris, C. C. Visualization of the transparent, gelatinous house of the pelagic tunicate *Oikopleura vanhoeffeni* using sepia ink. *Biol. Bull.* **178**, 118–125 (1990).
37. Conley, K. R. & Sutherland, K. R. Particle shape impacts export and fate in the ocean through interactions with the globally abundant appendicularian *Oikopleura dioica*. *PLoS ONE* **12**, e0183105 (2017).
38. Alldredge, A. L. Field behavior and adaptive strategies of appendicularians (Chordata: Tunicata). *Mar. Biol.* **38**, 29–39 (1976).

39. Flood, P. R. House formation and feeding behaviour of *Fritillaria borealis* (Appendicularia: Tunicata). *Mar. Biol*. **143**, 467–475 (2003).
40. Fenaux, R. Rhythm of secretion of Oikopleurid's houses. *Bull. Mar. Sci*. **37**, 498–503 (1985).
41. Purcell, J. E., Sturdevant, M. V. & Galt, C. P. in *Response of Marine Ecosystems to Global Change: Impact of Appendicularians* (eds Gorsky, G. et al.) 359–435 (Contemporary Publishing International, 2005).
42. Engelmann, J., Hanke, W., Mogdans, J. & Bleckmann, H. Hydrodynamic stimuli and the fish lateral line. *Nature* **408**, 51–52 (2000).
43. Janssen, J. & Strickler, J. R. in *Communication in Fishes*, vol. 1 (eds Ladich, F. et al.) 207–222 (Science Publishers, 2006).
44. Montgomery, J. C. in *The Mechanosensory Lateral Line* 561–574 (Springer, 1989).
45. Batchelor, G. K. *An Introduction to Fluid Dynamics* (Cambridge Univ. Press, 1967).

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Code availability

The custom MATLAB code developed as part of this study can be downloaded from our public repository at https://bitbucket.org/mbari/batho3dr.

Data availability

The data reported in this paper are archived and can be openly accessed using MBARI's Video Annotation and Reference System (VARS) query tool (https://www.mbari.org/products/research-software/video-annotation-and-reference-system-vars/query-interface/) with the search term 'Nature20190609559'. In addition, the data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgements We thank D. Graves, C. Kecy, D. Klimov, J. Erickson and MBARI technical staff for their engineering contributions to the development of DeepPIV, the crews of RVs *Rachel Carson* and *Western Flyer*, and the pilots of ROVs *Doc Ricketts*, *Ventana* and *MiniROV* for their contributions to this project. This work is a contribution of the Deep Ocean Inspiration Group and was supported by the David and Lucile Packard Foundation.

Author contributions K.K., B.H.R. and A.D.S. devised the experiments; K.K., G.T. and J.D. conducted the visualizations and processed the data; K.K., J.D., K.L. and R.E.S. analysed the data; K.K., B.H.R. and R.E.S. wrote the manuscript; all authors edited the manuscript.

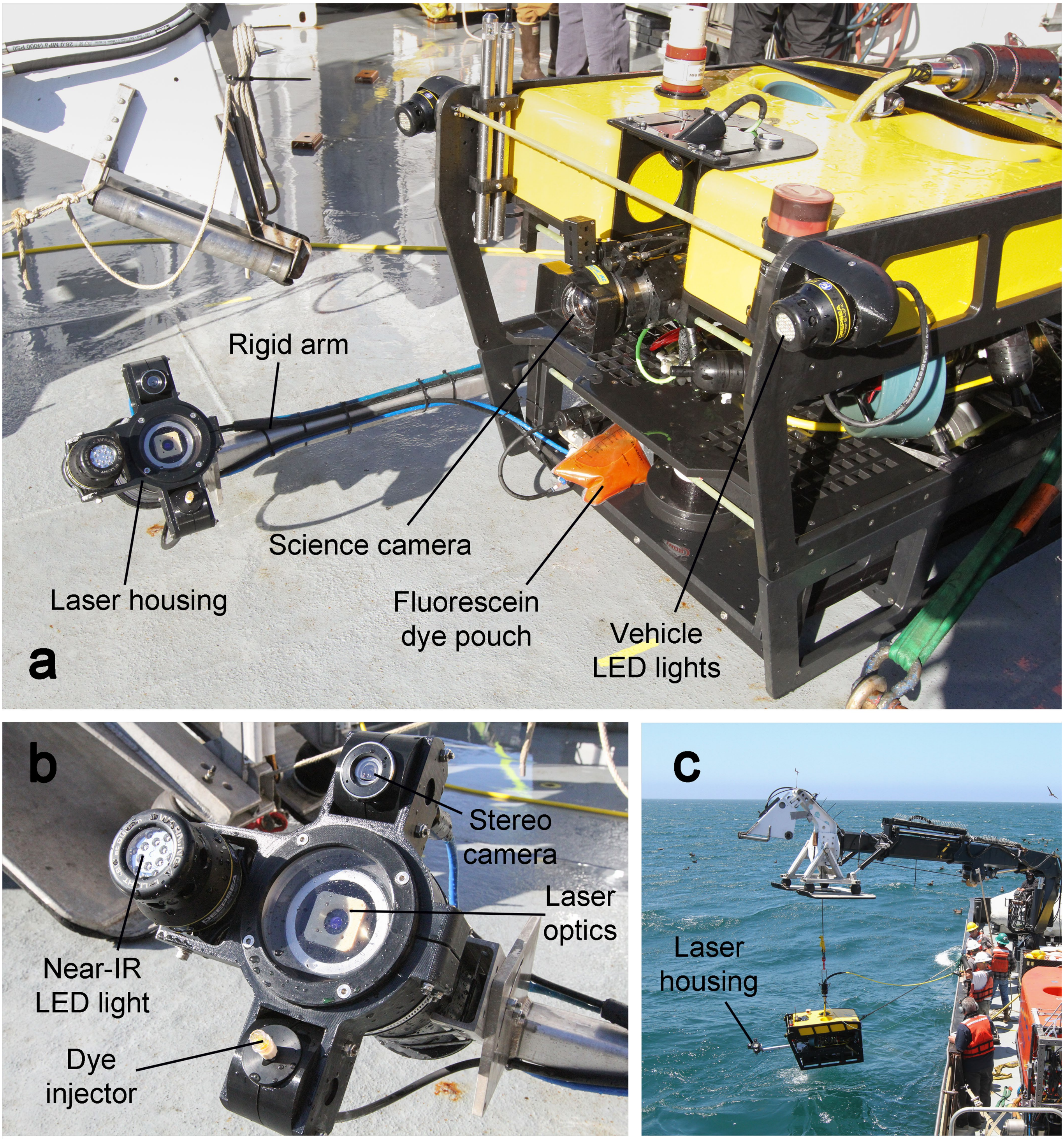
Competing interests The authors declare no competing interests.

Additional information

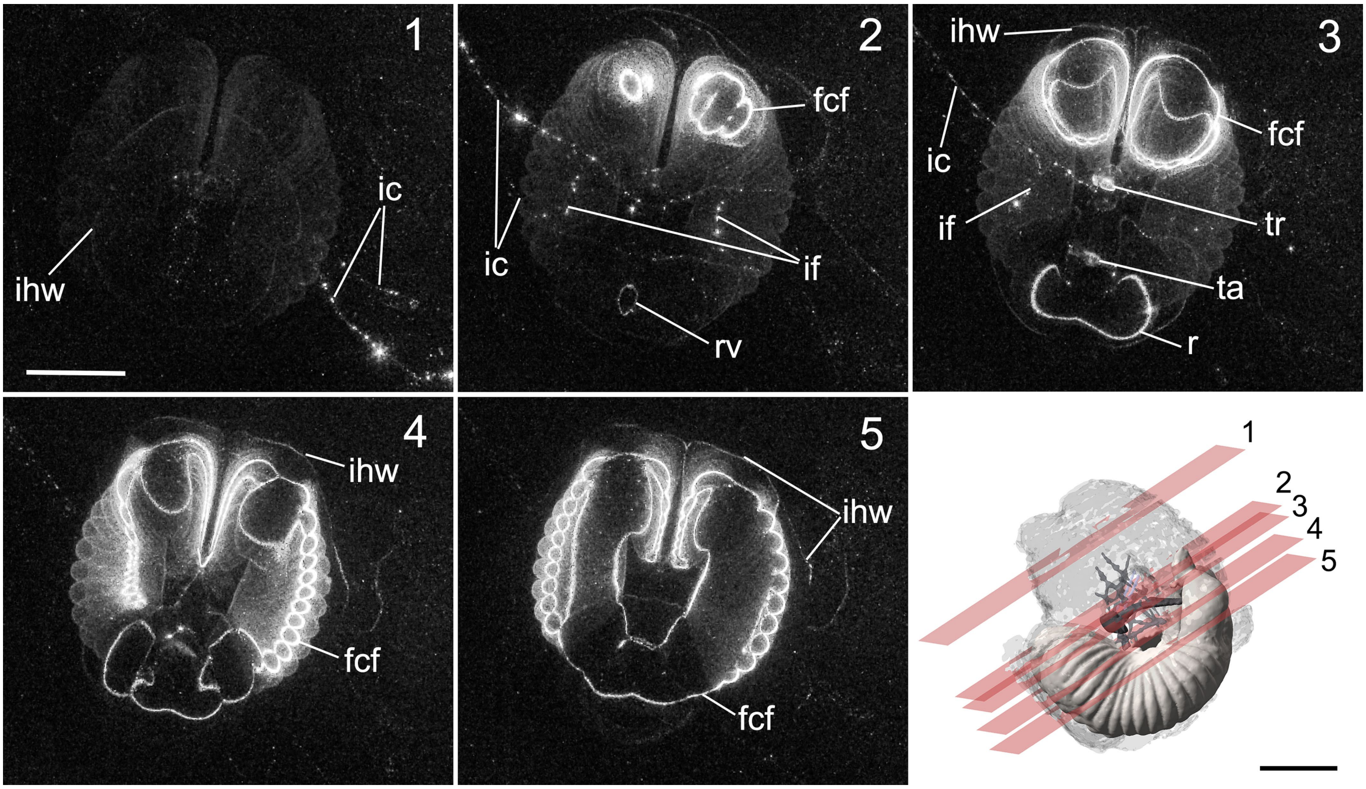
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2345-2.

Correspondence and requests for materials should be addressed to K.K.

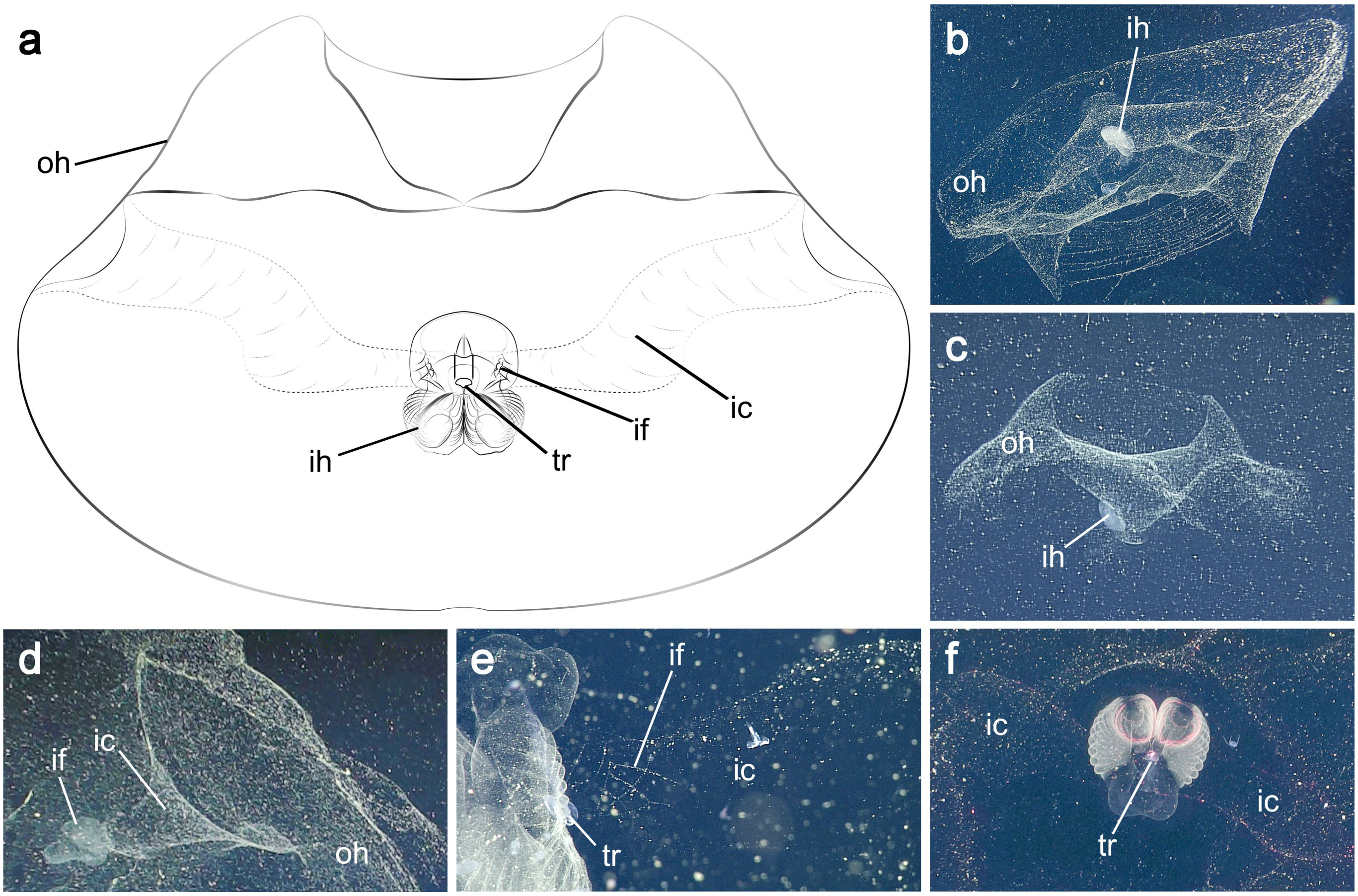
Peer review information *Nature* thanks Cornelia Jaspers, Kelly Sutherland and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



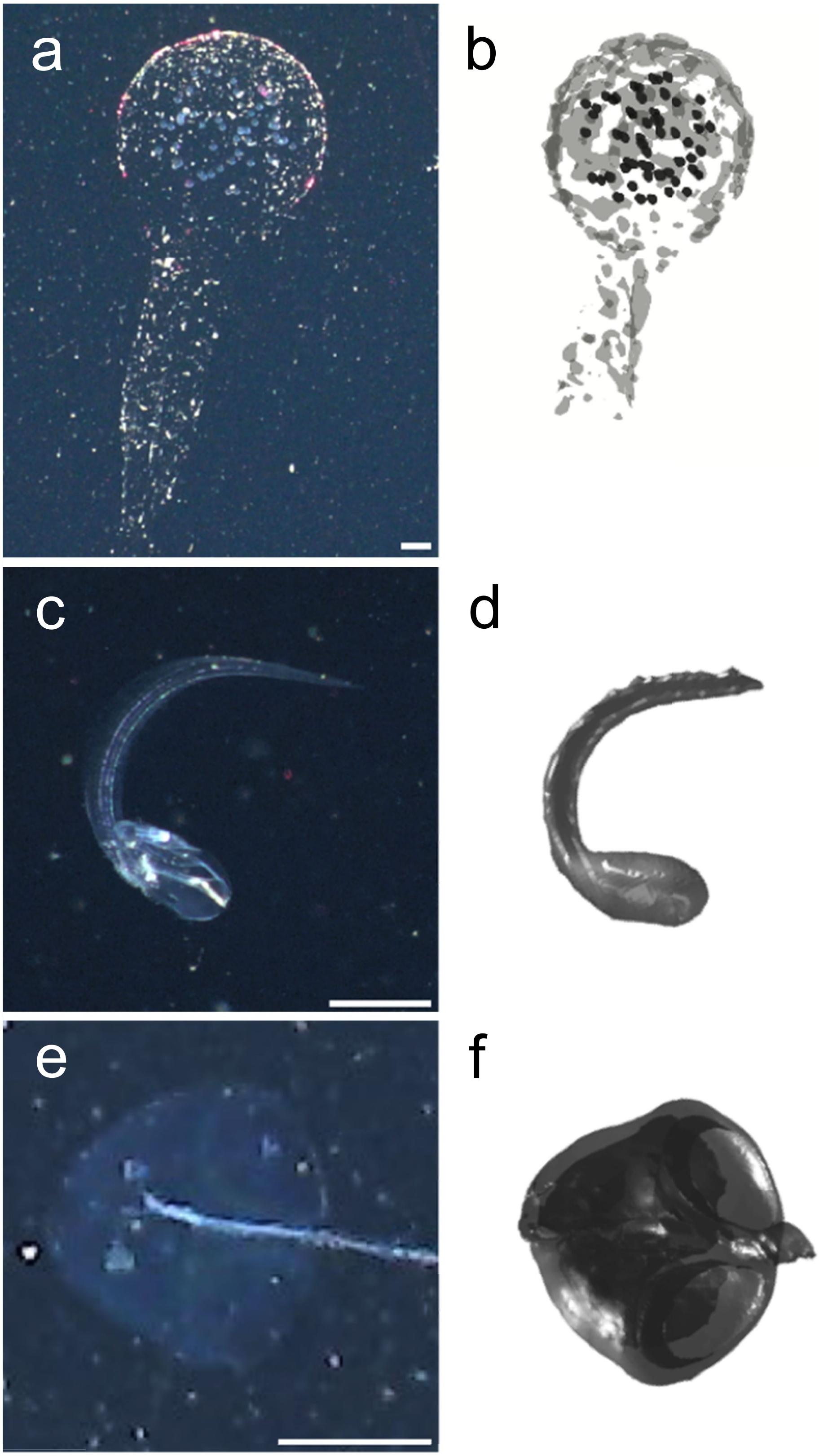
Extended Data Fig. 1 | DeepPIV hardware and deployment. a, DeepPIV is used to visualize gelatinous or mucus structures and conduct in situ three-dimensional scanning laser reconstructions using ROV *MiniROV*. **b**, Enlarged view of DeepPIV components affixed to the laser housing to generate a laser-sheet and fluorescent-dye field, as well as components to aid in pilot control of the vehicle during ROV deployments. **c**, *MiniROV* being launched in Monterey Bay from RV *Rachel Carson*.



Extended Data Fig. 2 | DeepPIV scans yield cross-sectional structural information. During a single laser sheet scan using DeepPIV, multiple planes (1–5 from dorsal to ventral) are illuminated to reveal different features in the mucus house structure of *B. stygius*. Scale bars, 4 cm.



Extended Data Fig. 3 | The inner and outer house as well as the connective mucus structures of *B. stygius*. **a**, Line drawing of the typical structure of the outer house and inlet channels with embedded inlet filters near the animal trunk. **b**, **c**, Overviews of the outer house structure with the animal–house complex oriented downwards (**b**) and upwards (**c**). **d**–**f**, Magnified views of the two inlet channels connecting laterally to the inner house from outside the outer house looking laterally (**d**), inside the inlet channel looking laterally (**e**) and inside the outer house looking dorsally (**f**).



Extended Data Fig. 4 | Three-dimensional reconstructions of mucus and gelatinous structures using DeepPIV. a–**f**, White-light illumination (**a**, **c**, **e**) provides two-dimensional snapshots of structures in midwater, where the scanning laser illumination of DeepPIV (**b**, **d**, **f**) can yield three-dimensional reconstructions of floating egg masses (**a**, **b**), larvacean bodies (**c**, **d**) and other gelatinous or mucus structures such as siphonophore swimming bells (**e**, **f**; *Desmophyes annectens*). Scale bars, 1 cm.

nature research

Corresponding author(s): Kakani Katija

Last updated by author(s): Jan 16, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a | Confirmed

☒ ☐ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement

☒ ☐ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly

☒ ☐ The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.

☒ ☐ A description of all covariates tested

☒ ☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons

☒ ☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)

☒ ☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted
Give P values as exact values whenever suitable.

☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings

☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes

☒ ☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used.
Data analysis	MATLAB R2017b (The Mathworks, Inc., USA) was used to develop and run the custom code for this project. The code can be downloaded from our public repository at https://bitbucket.org/mbari/batho3dr . Meshlab v2016.12 (P. Cignoni et al., Italy) was used to smooth the models. Adobe Photoshop v20.0 (Adobe, USA) was used for visualization and annotations of 3D reconstructions and assembly of figures. Adobe Premiere Pro v13.1 was used to create the supplementary videos.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data reported in this paper are archived, and can be openly accessed by using MBARI’s VARS Query tool (https://www.mbari.org/products/research-software/video-annotation-and-reference-system-vars/query-interface/) with the search term “Nature20190609559”.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We developed novel, in situ laser-imaging technology to reconstruct 3D models of giant larvacean mucus house structures. The models provide unprecedented views of once-enigmatic midwater structures, and elucidate the role that house structure plays in food capture and predator avoidance.
Research sample	We specifically targeted the species Bathochordaeus stygius between 100 and 300 m depths in the Monterey Bay National Marine Sanctuary. Samples were chosen based on animal size (>0.5 cm trunk length) and condition of the mucus structure (e.g., fully expanded or decrepit). Animal size was an important parameter, as fine control of the remotely operated vehicle during the scans became more challenging as animal size decreased. The size range between 1-2 cm trunk lengths represent fully-grown, adult individuals, and our 3D reconstructions are representative of that size range.
Sampling strategy	During deployments between 100 and 400 m depths, 71 specimens of the genus Bathochordaeus were observed. Samples were chosen based on animal size (>0.5 cm trunk length) and condition of the mucus structure (e.g., fully expanded or decrepit). Of these, three-dimensional reconstruction (or 3DR) measurements were conducted on 14 individual B. stygius houses. After using strict criteria for 3DR data selection (see data exclusions), 6 in situ reconstruction data sets corresponding to 5 individuals ranging in trunk length from 0.4 to 2.1 cm were used for subsequent analysis. These samples are representative of the size classes we commonly see of giant larvaceans in Monterey Bay.
Data collection	DeepPIV was deployed on 13 separate dives affixed to MBARI's MiniROV, a fly-away vehicle that is rated to 1500 m. Video from the DeepPIV instrument were recorded by the researchers on the research vessel, and the data are maintained in the VARS database.
Timing and spatial scale	The timing and scale were set by the frequency and location of pre-planned research expeditions. 13 ROV dives were made on 11 different days during 5 different expeditions in 2015. 2 dives were made in June, one in July, 4 in August on a single expedition, two in November, and 4 on a single expedition in December. Sampling frequency was constrained both spatially as well as temporally by other research activities performed during the same expeditions. Time of day was determined by operational requirements and ROV crew. All sampling was done in the Monterey Bay along the Monterey Canyon, at sites where water depth exceeds 1000 meters. Data provided here were obtained at sites less than 20 km apart.
Data exclusions	Some data were excluded based on pre-established criteria to limit any artifacts due to the nature of in situ image sampling using a remotely operated vehicle. Criteria for selecting clips for 3DR include (1) translation in the fore/aft direction of the vehicle with minimal rotation or translation in other axes, (2) vehicle motion at a nearly constant speed (< 10 cm/s), and (3) conditions (1) and (2) are met for a distance equal to the extent of the target. In addition, data were excluded if no positive species ID could be obtained.
Reproducibility	Multiple 3DR scans were conducted per individual observed, and with multiple individuals to yield models of larvacean house structure as shown here.
Randomization	Given the challenges of finding targets in midwater during limited cruise durations and robot deployments, randomization was not used.
Blinding	Given the challenges of finding targets in midwater during limited cruise durations and robot deployments, blinding was not used.

Did the study involve field work? ☒ Yes ☐ No

Field work, collection and transport

Field conditions	Field work were conducted during the daytime during all ROV dives in Monterey Bay.
Location	Dives were conducted between 100 and 400 m depths in Monterey Bay from June to December, 2015.
Access and import/export	Observations were made on board RVs Western Flyer and Rachel Carson in the Monterey Bay National Marine Sanctuary (MBNMS). Activities were conducted under the MBARI institutional permit with MBNMS and CA-DFW Scientific Collecting Permit #13337.
Disturbance	No disturbance was made as part of this field study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Not applicable.
Wild animals	Animals were observed non-invasively in the wild. We observed giant larvaceans, Bathochordaeus stygius and B. mcnutti, during this study.
Field-collected samples	This study did not involve animal samples collected from the field. Only in situ video data were obtained for this study.
Ethics oversight	No ethical approval was required as the study was non-invasively conducted in situ on invertebrates.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Mapping and characterization of structural variation in 17,795 human genomes

https://doi.org/10.1038/s41586-020-2371-0

Received: 29 December 2018

Accepted: 18 May 2020

Haley J. Abel[1,2,68], David E. Larson[1,2,68], Allison A. Regier[1,3], Colby Chiang[1], Indraniel Das[1], Krishna L. Kanchi[1], Ryan M. Layer[4,5], Benjamin M. Neale[6,7,8], William J. Salerno[9], Catherine Reeves[10], Steven Buyske[11], NHGRI Centers for Common Disease Genomics*, Tara C. Matise[12], Donna M. Muzny[9], Michael C. Zody[10], Eric S. Lander[6,13,14], Susan K. Dutcher[1,2], Nathan O. Stitziel[1,2,3] & Ira M. Hall[1,2,3]✉

A key goal of whole-genome sequencing for studies of human genetics is to interrogate all forms of variation, including single-nucleotide variants, small insertion or deletion (indel) variants and structural variants. However, tools and resources for the study of structural variants have lagged behind those for smaller variants. Here we used a scalable pipeline[1] to map and characterize structural variants in 17,795 deeply sequenced human genomes. We publicly release site-frequency data to create the largest, to our knowledge, whole-genome-sequencing-based structural variant resource so far. On average, individuals carry 2.9 rare structural variants that alter coding regions; these variants affect the dosage or structure of 4.2 genes and account for 4.0–11.2% of rare high-impact coding alleles. Using a computational model, we estimate that structural variants account for 17.2% of rare alleles genome-wide, with predicted deleterious effects that are equivalent to loss-of-function coding alleles; approximately 90% of such structural variants are noncoding deletions (mean 19.1 per genome). We report 158,991 ultra-rare structural variants and show that 2% of individuals carry ultra-rare megabase-scale structural variants, nearly half of which are balanced or complex rearrangements. Finally, we infer the dosage sensitivity of genes and noncoding elements, and reveal trends that relate to element class and conservation. This work will help to guide the analysis and interpretation of structural variants in the era of whole-genome sequencing.

Human genetics studies use whole-genome sequencing (WGS) to enable comprehensive trait-mapping analyses across the full diversity of genome variation, including structural variants (SVs) of 50 base pairs (bp) or greater, such as deletions, duplications, insertions, inversions and other rearrangements. Previous work suggests that SVs have a disproportionately large role (relative to their abundance) in the biology of rare diseases[2] and in shaping heritable differences in gene expression in the human population[3–5]. Rare and de novo SVs have been implicated in the genetics of autism[6–10] and schizophrenia[11–14], but few other complex trait association studies have directly assessed SVs[15,16].

One challenge for the interpretation of SVs in WGS-based studies is the lack of high-quality publicly available variant maps from large populations. Our current knowledge is based primarily on three sources: (1) a large and disparate collection of array-based studies[17–19], with limited allele-frequency data and low resolution; (2) the 1000 Genomes Project callset[5], which has been invaluable but is limited by the modest sample size and low-coverage design; and (3) an assortment of smaller WGS-based studies with varied coverage, technologies, methods of analysis and levels of data accessibility[8,9,20–22].

There is an opportunity to improve our knowledge of SVs in human populations through the systematic analysis of large-scale WGS data resources that are generated by initiatives such as the National Human Genome Research Institute (NHGRI) Centers for Common Disease Genomics (CCDG). A key barrier to the creation of larger and more-informative catalogues of SVs is the lack of computational tools that can scale to the size of ever-growing datasets. To this aim, we have developed an SV analysis pipeline that is open source and highly scalable[1], and used it to map and characterize SVs in 17,795 deeply sequenced human genomes.

A population-scale map of SVs

The samples analysed here are derived from case–control studies and quantitative trait-mapping collections of common diseases that

---

[1]McDonnell Genome Institute, Washington University School of Medicine, St Louis, MO, USA. [2]Department of Genetics, Washington University School of Medicine, St Louis, MO, USA. [3]Department of Medicine, Washington University School of Medicine, St Louis, MO, USA. [4]BioFrontiers Institute, University of Colorado, Boulder, CO, USA. [5]Department of Computer Science, University of Colorado, Boulder, CO, USA. [6]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [7]Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA. [8]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, USA. [9]Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA. [10]New York Genome Center, New York, NY, USA. [11]Department of Statistics, Rutgers University, Piscataway, NJ, USA. [12]Department of Genetics, Rutgers University, Piscataway, NJ, USA. [13]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA. [14]Department of Systems Biology, Harvard Medical School, Boston, MA, USA. [68]These authors contributed equally: Haley J. Abel, David E. Larson. *A list of authors and their affiliations appears at the end of the paper. ✉e-mail: ihall@genome.wustl.edu

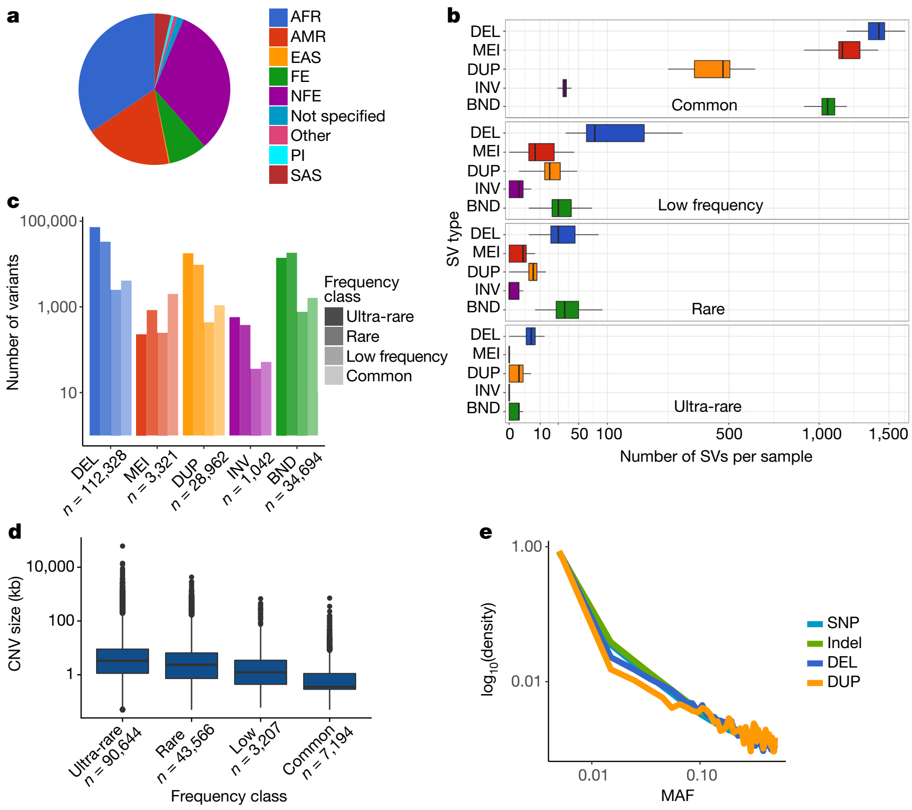


Fig. 1 | The public version of the B38 callset derived from 14,623 samples. **a**, Self-reported ancestry. AFR, African; AMR, admixed American; EAS, East Asian; FE, Finnish European; NFE, non-Finnish European; PI, Pacific Islander; SAS, South Asian. **b**, Number of SVs per sample (*x* axis, square-root-scaled) by SV type (*y* axis) and frequency class. SV types are: deletion (DEL), mobile-element insertion (MEI), duplication (DUP), inversion (INV) and breakend (BND). MAF bins are defined as ultra-rare (unique to an individual or family), rare (MAF < 1%), low frequency (1% < MAF < 5%) or common (MAF > 5%). **c**, Number of high-confidence SVs by type and frequency bin. **d**, CNV length distributions for each frequency class. **e**, MAF distribution for SNV (*n* = 85,687,916), indel (*n* = 9,477,540), DEL (*n* = 43,872) and DUP (*n* = 10,805) variants for a subset of 4,298 samples for which Genome Analysis Toolkit (GATK)-based SNV and indel calls were also available. All box plots in this figure indicate the median (centre line) and the first and third quartiles (box limits). The upper whiskers extend to the lesser extreme of the maximum and the third quartile plus 1.5 times the interquartile range (IQR); the lower whiskers extend to the lesser extreme of the minimum and the first quartile minus 1.5 times the IQR.

were sequenced under the CCDG programme, supplemented with ancestrally diverse samples from the Population Architecture Using Genomics and Epidemiology (PAGE) consortium and the Simons Genome Diversity Panel. The final ancestry composition includes 24% African, 16% Latino, 11% Finnish, 39% non-Finnish European and 9% other diverse samples from around the world (Extended Data Table 1).

The tools and pipelines used for this work are described elsewhere[1]. In brief, we developed a highly scalable software toolkit (svtools) and workflow for the generation of SV callsets on a large scale, which combines per-sample variant discovery[23], resolution-aware cross-sample merging, breakpoint genotyping[24], copy-number annotation and variant classification (Extended Data Fig. 1). We created two distinct SV callsets using different reference genome and pipeline versions. The ‘B37’ callset includes 118,973 high-confidence SVs from 8,426 samples that were sequenced at the McDonnell Genome Institute and aligned to the GRCh37 reference genome. The ‘B38’ callset includes 241,031 high-confidence SVs from 23,175 samples that were sequenced at four CCDG sites and aligned to GRCh38 using the ‘functional equivalence’ pipeline[25] (Methods). Of the 26,347 distinct samples in the union of the two callsets, aggregate-level sharing is permitted for 17,795; these make up the official public release (Supplementary Files 1, 2). For simplicity of presentation, most analyses below focus on the larger B38 callset (Supplementary Table 1).

We observed a mean of 4,442 high-confidence SVs per genome—predominantly deletions (35%), mobile-element insertions (MEIs) (27%) and tandem duplications (11%) (Fig. 1b, Extended Data Figs. 2, 3). Variant counts and linkage disequilibrium patterns are consistent with previous studies that used similar methods[4,5], and most SVs are mapped to base-pair resolution (Extended Data Figs. 2, 3). As expected, the site-frequency spectrum approximates that of single-nucleotide variants (SNVs) and indels, the size distribution shows increasing length with decreasing frequency, and principal component analysis (PCA) reveals a population structure that is consistent with self-reported ancestry (Fig. 1, Extended Data Figs. 2–4). Per-genome SV counts are broadly consistent and vary as expected on the basis of ancestry, with more genetic variation in individuals of African ancestry and fewer singletons in Finnish individuals (Extended Data Figs. 2, 3). Although we observe some technical variability owing to cohort and sequencing centre, these effects are mainly limited to small (less than 1 kb) copy-number variants (CNVs) that are detected solely by read-pair signals, which are sensitive to methods of library preparation and alignment filtering (Methods, Extended Data Fig. 3).

We further characterized callset quality using independent data and analyses (Supplementary Note) including (1) validation by deep-coverage (greater than 50×) long-read data from nine genomes; (2) sensitivity relative to a comprehensive long-read callset[26]; (3) inheritance patterns within a set of three-generation pedigrees; and (4) comparison to well-characterized short-read callsets[5,26] (Supplementary Tables 2–4, Extended Data Figs. 5–7). We achieve a validation rate of 84% by long-read data, with higher validation rates for the variant

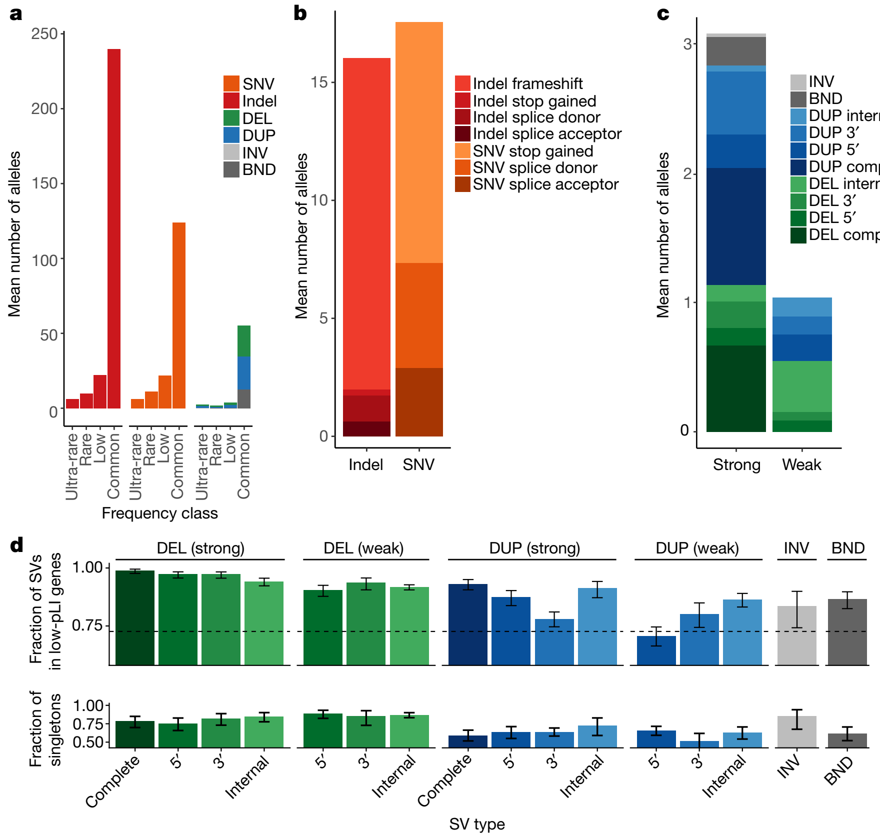


Fig. 2 | Burden of rare gene-altering SVs. a, Mean number of gene alterations per sample by type and frequency class ($n = 4{,}298$ samples). **b**, Mean number of rare (MAF < 1%) high-confidence protein-truncating variants per sample by type and VEP consequence. **c**, Mean number of rare (MAF < 1%) SV-derived gene alterations per sample by type. DEL and DUP are classified into strong (affecting more than 20% of exons of the principal transcript) and weak (affecting less than 20% of exons of the principal transcript) and sub-classified as internal (variant overlaps at least one coding exon, but neither the 3′ nor the 5′ end of the principal transcript), 3′ (variant overlaps the 3′ end of the transcript), 5′ (variant overlaps the 5′ end of the transcript) and complete (variant overlaps all coding exons in the principal transcript). **d**, Top, fraction of rare (MAF < 1%) gene-altering variants occurring in genes with a low pLI score (pLI < 0.9) by SV type and size class, stratified by affected gene region in the B38 callset ($n = 14{,}623$). The dotted line indicates the expected fraction, assuming a uniform distribution of SVs in coding exons. Bottom, fraction of singletons for gene-altering variants by type in the B38 callset ($n = 14{,}623$), restricted to genes with pLI > 0.1. Error bars (**d**, **e**) indicate 95% confidence intervals (Wilson score method). See Supplementary Table 5 for the number of variants in each category.

classes that are most relevant to the findings below: deletions (87%), rare SVs (90%) and singleton SVs (95%). On the basis of the validation rates of SV frequency classes and their relative abundance in the full dataset, we estimate a false discovery rate of 7.0%. Although the overall sensitivity is low (49%) compared to long-read SV maps—owing to the inherent difficulty of detecting repetitive variants from short reads—it is comparable to published short-read callsets[4,5,26] and is substantially higher for functionally relevant subtypes, such as SVs larger than 1 kb (63%) and predicted high-impact variants (82%).

Burden of deleterious rare SVs

The contribution of rare SVs to human disease remains unclear. Well-powered WGS-based trait-mapping studies will ultimately be required to address this; however, the overall burden of predicted pathogenic mutations in the human population is informative and can be estimated from our data. Our analysis of 14,623 individuals identified 42,765 rare SV alleles (minor allele frequency (MAF) of less than 1%) that are predicted to decrease gene dosage ($n = 9{,}416$), alter gene function (for example, single exon deletion; $n = 26{,}337$) or increase gene dosage ($n = 7{,}012$). The majority of rare gene-altering SVs are deletions (54.5%), with fewer duplications (42.2%) and a small fraction of other variant types, primarily inversions and complex rearrangements that interrupt or rearrange exons. Of these, 23.4% affect multiple genes and 10.4% affect three or more genes, resulting in a mean of 4.2 SV-altered genes per individual. On the basis of a strict definition of loss-of-function SVs—gene disruptions and gene deletions that affect more than 20% of exons—we identified a mean of 1.39 rare SV-based loss-of-function alleles per person. An analysis of 4,298 samples with SV calls and SNV or indel calls revealed that individuals carry a mean of 33.6 rare high-confidence loss-of-function SNVs and small indels (Fig. 2), consistent with previous studies[27]. Thus, SVs account for 4–11.2% of rare, predicted high-impact gene alterations in a population sample, depending on whether we consider all coding SVs or a strictly defined set of loss-of-function variants (Fig. 2c). These are likely to be underestimates, considering that the false-negative rate of SV detection is typically higher than that of SNVs and small indels[24,26].

To characterize the relative effect of different coding SV classes we calculated two measures of purifying selection (Fig. 2d): (1) the fraction of variants that affect dosage-tolerant genes with a loss-of-function intolerance (pLI)[27,28] score of less than 0.9; and (2) the fraction of variants that are present as singletons found in only one individual or family. By these measures, deletions are more deleterious than duplications, and complete gene deletions are the most deleterious class. Notably, on the basis of the fraction of variants in dosage-intolerant genes, complete gene duplications and sub-genic deletions that affect fewer than 20% of exons are relatively depleted; this suggests that many gene-altering

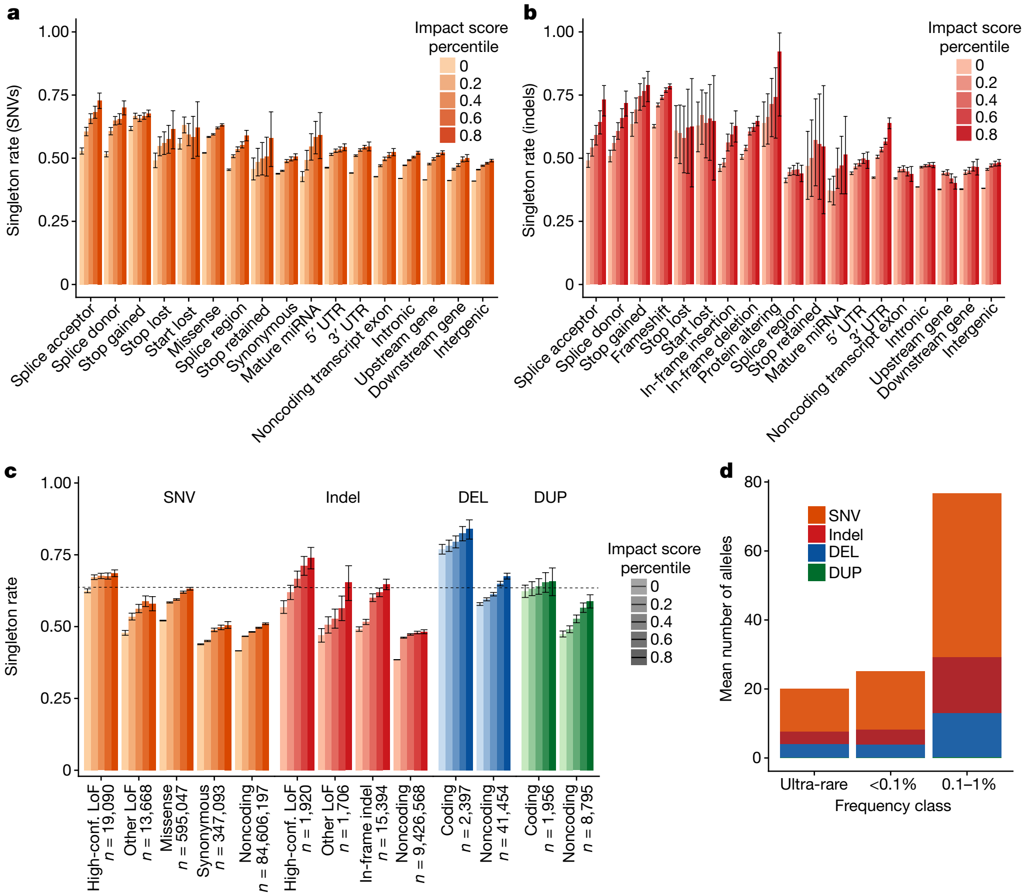


Fig. 3 | Estimation of genome-wide burden of high-impact functional alleles. **a**, Singleton rates for SNVs, by VEP consequence and percentile of the impact score (derived from combined LINSIGHT and CADD impact scores). miRNA, microRNA; UTR, untranslated region. **b**, Singleton rates for indels. **c**, Singleton rates by variant type and percentile of combined CADD–LINSIGHT impact score. The horizontal dotted line shows the singleton rate for all high-confidence (high-conf.) SNV or indel loss-of-function (LoF) mutations. 'Other LoF' indicates VEP-annotated protein-truncating variants that are not classified as high confidence by LOFTEE. DELs and DUPs that intersect with any coding exon of the principal transcript are classified as coding; otherwise, they are noncoding. **d**, Mean number of strongly deleterious alleles genome-wide per sample, by type and frequency class. Error bars (**a**–**c**) indicate 95% confidence intervals (Wilson score method). See Supplementary Table 6 for counts of variants in each category.

SVs are strongly deleterious, even those not predicted to completely obliterate gene function.

The above calculations ignore missense and noncoding variants, which are expected to make up a large fraction of rare functional variation. Predicting the effect of these variant types is challenging, but we can approximate their relative contribution to the deleterious variant burden under two simplifying assumptions: (1) impact-prediction algorithms such as CADD[29] and LINSIGHT[30] are capable of ranking variants within a given class (SNV, indel, SV) by their degree of deleteriousness; and (2) the mean deleterious impact of a given set of variants is reflected by its singleton rate. The first assumption is somewhat tenuous, but should be valid here given that impact-prediction inaccuracies are likely to affect all variant classes similarly; the second should hold under an infinite sites model of mutation, which is reasonable for the sample size (*n* = 4,298 samples) used in this analysis. We note that other evolutionary forces such as positive selection, background selection and biased gene conversion can also shape the site-frequency spectrum; however, we expect that these forces would act similarly on the variant classes examined here, given that this is a genome-wide analysis of a very large number of sites.

We used CADD and LINSIGHT to generate impact scores for SNVs, indels, deletions and duplications (Methods). As expected, these are highly correlated with singleton rate and variant effect predictions from the Ensembl Variant Effect Predictor (VEP)[31] and LOFTEE[27] (Fig. 3). We sought to identify 'strongly deleterious' variants from each class by choosing impact-score thresholds to match the singleton rate of the entire set of high-confidence loss-of-function mutations. Individuals carried a mean of 121.9 strongly deleterious rare variants, comprising 63% SNVs, 19.8% indels and 17.2% SVs (Fig. 3d). Given the relative numerical abundance of different rare variant classes, this suggests that a given rare SV is 841-fold more likely to be strongly deleterious than a rare SNV, and 341-fold more likely than a rare indel. Predicted deleterious SVs are slightly larger than rare SVs on the whole (median 4.5 versus 2.8 kb). Whereas only a minority (13.1%) of predicted strongly deleterious SNVs and indels are noncoding, 90.1% of predicted strongly deleterious rare SVs are noncoding. In particular, the top 50% of noncoding deletions show similar levels of purifying selection (as measured by singleton rate) as high-confidence loss-of-function variants that are caused by SNVs or indels (Fig. 3c), suggesting that a typical individual carries 19.1 alleles for strongly deleterious rare noncoding deletions. This suggests that noncoding deletions may have strongly deleterious effects, and may have a larger than expected role in human disease.

Landscape of ultra-rare SVs

Most ultra-rare SVs represent recent or de novo structural mutations, and thus the relative abundance of different classes of ultra-rare SV sheds light on the underlying mutational processes. We identified 158,991 ultra-rare SVs (105,175 high-confidence) that were

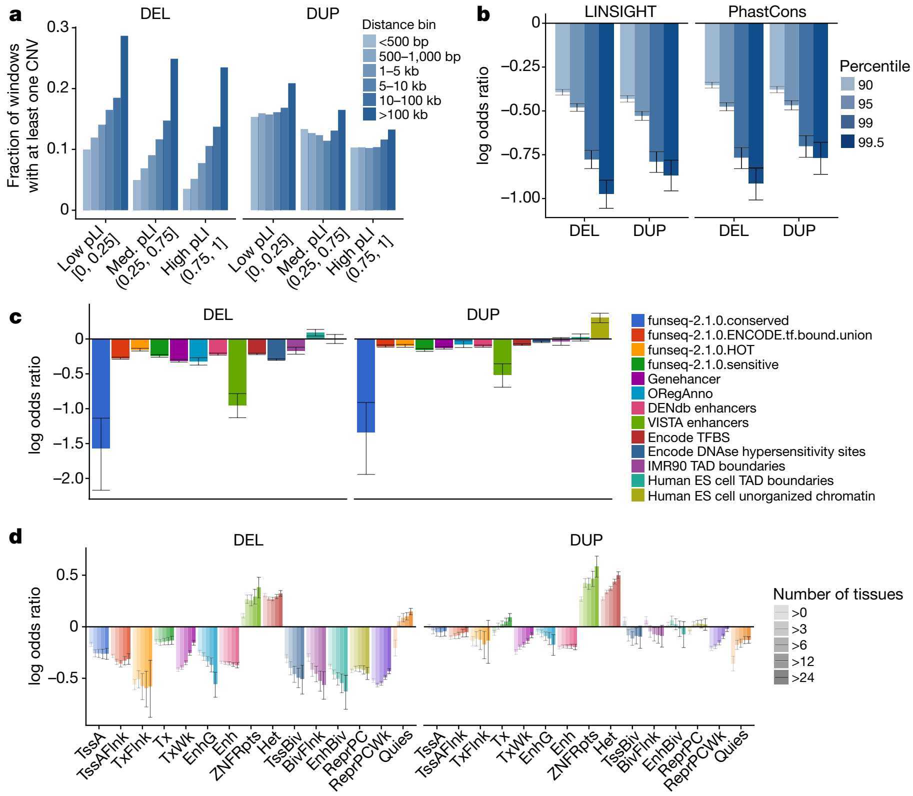


Fig. 4 | Dosage sensitivity of functional annotations. a, Fraction of 1-kb genomic windows that contain at least one CNV, as a function of the distance to the nearest coding exon and the pLI of that gene. **b**, Depletion of CNVs in conserved genomic regions. Odds ratios (log-transformed) for the occurrence of CNVs in highly conserved (based on LINSIGHT or PhastCons percentile) versus less-conserved regions. Odds ratios are Cochran–Mantel–Haenszel estimates, stratified by the distance to the nearest coding exon and the pLI of that gene. **c**, Odds ratios (log-transformed; estimated as in **b**) for the occurrence of CNVs in 1-kb windows that intersect various functional annotation tracks. Human ES cell, human embryonic stem cell; TAD, topologically associated domain; TFBS, transcription-factor-binding site. **d**, Odds ratios (log-transformed; estimated as in **b**) for the occurrence of CNVs in 1-kb windows that overlap Roadmap segmentations, stratified by the number of Roadmap tissues in which the region is observed. Error bars (**b**–**d**) indicate 95% confidence intervals estimated by block bootstrap.

present in only one of 14,623 individuals or were unique to a family. This corresponds to a mean of around 11.4 per individual (Extended Data Fig. 8a). Ultra-rare SVs are mainly composed of deletions (5.2 per person) and duplications (1.3), with a smaller number of inversions (0.17).

It is notable that around 40% of ultra-rare SV breakpoints in our dataset cannot be readily classified into the canonical forms of SV. This is a known limitation of short-read WGS, and such variants are often ignored. Formally, these SVs are of the 'breakend' (BND) class, which is a generic term in the VCF specification for SV breakpoints that cannot be unequivocally classified[32]. We examined the 63,559 ultra-rare BNDs for insights into their composition and origin. Many (17.0%) appear to be deletions that are too small (less than 100 bp) to exhibit convincing read-depth support, and that our pipeline conservatively classifies as BNDs (for example, complex SVs can masquerade as deletions). Some (2.4%) of the ultra-rare BNDs stem from 1,542 'retrogene insertions', which are caused by retroelement machinery acting on mRNAs. This set of retrogene insertions is around 10-fold larger than those of previous maps[33–35] and will be valuable for future studies. Another 5.5% of ultra-rare BNDs are complex genomic rearrangements with multiple breakpoints in close proximity (less than 100 kb). The remainder are variants that are difficult to classify, which involve local (49.9%, within 1 Mb) or distant (5.7%, more than 1 Mb apart) intra-chromosomal alterations or inter-chromosomal alterations (27.2%), and of which many (78.0%) are classified as low-confidence SV calls. This final class is probably caused primarily by variation in repetitive elements, but is also expected to be enriched for false positives.

A variety of sporadic disorders are caused by extremely large and/or complex SVs, but—owing to the limitations of the array-based methods that have been used in previous large-scale studies[36], which fail to detect balanced events or resolve complex variant architectures—our knowledge of the frequency and architecture of these marked alterations in the general population is incomplete. We observed 138 megabase-scale CNVs, which corresponds to a frequency of around 0.01 per individual; these include 47 deletions and 91 duplications, and affect a mean of 12.1 genes (Extended Data Fig. 8b). Three individuals carried two megabase-scale CNVs, apparently owing to independent mutations. We observed 19 reciprocal translocations (0.001 per individual), consistent with previous cytogenetic-based estimates[37,38]. Of these translocations, 14 affect one gene and two affect two genes, producing one predicted in-frame gene fusion (PI4KA:MGLL). We applied breakpoint clustering (as in a previous study[39]) to identify ultra-rare complex rearrangements and discovered 33 complex SVs that span more than 1 Mb (0.003 per individual). Most of these (20 out of 33, 60.6%) involve three breakpoints; however, we observed five large-scale rearrangements with five or more breakpoints. Notably, when the entire SV size distribution is considered, 3.3% of ultra-rare SVs are complex variants, which is consistent with previous smaller-scale studies[40–44].

Dosage sensitivity

A motivation for creating population-scale SV maps is to annotate genomic regions on the basis of their tolerance to dosage changes and structural rearrangements, thus revealing the genes and noncoding elements that are most important (or dispensable) for human development and viability. The pLI score from the Exome Aggregation Consortium (ExAC) and the Genome Aggregation Database (gnomAD)[27,28] has proven invaluable for this purpose, but does not predict the effects of increased dosage or include noncoding elements.

We first generated deletion (DEL) and duplication (DUP) sensitivity scores for each gene on the basis of the observed frequency of CNVs in the combined dataset of 17,795 samples (as in a previous study[45]; see Methods). The resulting scores correlate with the CNV scores from ExAC[45], and with the DECIPHER haploinsufficiency score[46] (Extended Data Fig. 9). Despite their relatively modest correlations with one another, all three measures are informative compared with pLI, which was generated using an independent set of variants (SNVs and indels). A combined score from multiple datasets performs better than any single score, and may be useful for interpreting rare SVs (Supplementary File 4).

We next performed a genome-wide analysis based on the frequency of dosage alterations in 1-kb genomic windows (Methods). Our current dataset is not large enough to predict dosage-sensitive noncoding elements on the basis of the absence of variation; however, we can investigate the relative sensitivity of genomic features in aggregate. As expected, we observed a strong depletion of CNVs near coding exons, which varied according to the proximity to the nearest exon as well as the pLI of the corresponding gene (Fig. 4a). We therefore estimated odds ratios for depletion of CNVs in each functionally annotated region, stratified by distance to and pLI of the nearest exon. The resulting dosage-sensitivity scores mirror independent measures of selective constraint including LINSIGHT and PhastCons (Fig. 4b).

We also examined the relative dosage sensitivity of regulatory and epigenomic annotations from various projects[47–52] (Fig. 4). Regulatory elements such as enhancers, polycomb repressors, DNase hypersensitivity sites and transcription-factor-binding sites show strong sensitivity to dosage loss through deletion, whereas regions of inert noncoding annotations do not. The patterns of sensitivity to dosage gains through duplication are broadly similar, albeit weaker, with no obviously distinct patterns at (for example) enhancers, repressors or insulators. The dosage sensitivity of regulatory elements at ‘bivalent’ genomic regions from the NIH Roadmap Epigenomics project is greater than their counterparts (for example, enhancers versus bivalent enhancers), suggesting that such elements may be under especially strong selection. Furthermore, dosage sensitivity increases with the number of cell types that share a given annotation, suggesting that sensitivity is higher for constitutive regulatory elements compared to those that act in a more cell-type-specific manner.

Discussion

Here, we have conducted the largest—to our knowledge—WGS-based study of SVs in the human population so far. The sample size and use of deep (greater than 20×) WGS allowed us to map rare SVs at high genomic resolution and estimate the relative burden of deleterious SVs. Our data suggest that rare SVs account for 4–11.2% of deleterious coding alleles and 17.2% of deleterious alleles genome-wide—a disproportionate contribution considering that SVs comprise roughly 0.1% of variants. The burden of rare, strongly deleterious noncoding deletions that is apparent in our dataset is notable: we estimate that a typical individual carries 19.1 rare noncoding deletions that exhibit levels of purifying selection similar to loss-of-function SNVs and indels (of which there are 33.6 per individual). These results indicate that comprehensive assessment of SVs will improve power in rare-variant association studies.

The public site-frequency maps reported here will also aid the interpretation of variants in smaller-scale WGS-based studies (for example, through look-ups of allele frequency), in particular as they were generated by a systematic joint analysis of large datasets from diverse populations (similar to ExAC and gnomAD[27]). One limitation is the high false-negative rate for repetitive SVs, including MEIs, short tandem repeats (STRs) and multi-allelic CNVs, owing to the limitations of algorithms that rely on unique short-read alignments. Whereas we have reported a mean of 4,442 SVs per genome, recent long-read analyses predict up to around 27,662 SVs per genome, including STRs and other highly repetitive elements[26]. Although the inherent limitations of short-read WGS cannot be overcome, this resource could be made more comprehensive in future work with specialized algorithms tailored to MEIs, STRs and multi-allelic CNVs.

Finally, we have mined this resource to assess the dosage sensitivity of genes and noncoding elements. For genes, our results complement existing estimates from exome-sequencing and microarray data; for noncoding elements, we observe strong correlations with measures of nucleotide conservation, purifying selection, activity of regulatory elements and cell-type specificity. Although our current sample size is insufficient to assess the dosage sensitivity of individual noncoding elements, this will become feasible as large-scale WGS resources from ongoing international programs become available.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2371-0.

1. Larson, D. E. et al. svtools: population-scale analysis of structural variation. *Bioinformatics* **35**, 4782–4787 (2019).
2. Weischenfeldt, J., Symmons, O., Spitz, F. & Korbel, J. O. Phenotypic impact of genomic structural variation: insights from and for human disease. *Nat. Rev. Genet.* **14**, 125–138 (2013).
3. Stranger, B. E. et al. Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* **315**, 848–853 (2007).
4. Chiang, C. et al. The impact of structural variation on human gene expression. *Nat. Genet.* **49**, 692–699 (2017).
5. Sudmant, P. H. et al. An integrated map of structural variation in 2,504 human genomes. *Nature* **526**, 75–81 (2015).
6. Sebat, J. et al. Strong association of de novo copy number mutations with autism. *Science* **316**, 445–449 (2007).
7. Weiss, L. A. et al. Association between microdeletion and microduplication at 16p11.2 and autism. *N. Engl. J. Med.* **358**, 667–675 (2008).
8. Turner, T. N. et al. Genomic patterns of de novo mutation in simplex autism. *Cell* **171**, 710–722 (2017).
9. Werling, D. M. et al. An analytical framework for whole-genome sequence association studies and its implications for autism spectrum disorder. *Nat. Genet.* **50**, 727–736 (2018).
10. Brandler, W. M. et al. Paternally inherited cis-regulatory structural variants are associated with autism. *Science* **360**, 327–331 (2018).
11. International Schizophrenia Consortium. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* **455**, 237–241 (2008).
12. Walsh, T. et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* **320**, 539–543 (2008).
13. McCarthy, S. E. et al. Microduplications of 16p11.2 are associated with schizophrenia. *Nat. Genet.* **41**, 1223–1227 (2009).
14. Marshall, C. R. et al. Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. *Nat. Genet.* **49**, 27–35 (2017).
15. Craddock, N. et al. Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. *Nature* **464**, 713–720 (2010).
16. Kathiresan, S. et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat. Genet.* **41**, 334–341 (2009).
17. MacDonald, J. R., Ziman, R., Yuen, R. K., Feuk, L. & Scherer, S. W. The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res.* **42**, D986–D992 (2014).
18. Bragin, E. et al. DECIPHER: database for the interpretation of phenotype-linked plausibly pathogenic sequence and copy-number variation. *Nucleic Acids Res.* **42**, D993–D1000 (2014).
19. Lappalainen, I. et al. dbVar and DGVa: public archives for genomic structural variation. *Nucleic Acids Res.* **41**, D936–D941 (2013).
20. Hehir-Kwa, J. Y. et al. A high-quality human reference panel reveals the complexity and distribution of genomic structural variants. *Nat. Commun.* **7**, 12989 (2016).
21. Maretty, L. et al. Sequencing and de novo assembly of 150 genomes from Denmark as a population reference. *Nature* **548**, 87–91 (2017).

22. Sudmant, P. H. et al. Global diversity, population stratification, and selection of human copy-number variation. *Science* **349**, aab3761 (2015).
23. Layer, R. M., Chiang, C., Quinlan, A. R. & Hall, I. M. LUMPY: a probabilistic framework for structural variant discovery. *Genome Biol*. **15**, R84 (2014).
24. Chiang, C. et al. SpeedSeq: ultra-fast personal genome analysis and interpretation. *Nat. Methods* **12**, 966–968 (2015).
25. Regier, A. A. et al. Functional equivalence of genome sequencing analysis pipelines enables harmonized variant calling across human genetics projects. *Nat. Commun*. **9**, 4038 (2018).
26. Chaisson, M. J. P. et al. Multi-platform discovery of haplotype-resolved structural variation in human genomes. *Nat. Commun*. **10**, 1784 (2019).
27. Lek, M. et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).
28. 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
29. Kircher, M. et al. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet*. **46**, 310–315 (2014).
30. Huang, Y. F., Gulko, B. & Siepel, A. Fast, scalable prediction of deleterious noncoding variants from functional and population genomic data. *Nat. Genet*. **49**, 618–624 (2017).
31. McLaren, W. et al. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics* **26**, 2069–2070 (2010).
32. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
33. Ewing, A. D. et al. Retrotransposition of gene transcripts leads to structural variation in mammalian genomes. *Genome Biol*. **14**, R22 (2013).
34. Schrider, D. R. et al. Gene copy-number polymorphism caused by retrotransposition in humans. *PLoS Genet*. **9**, e1003242 (2013).
35. Abyzov, A. et al. Analysis of variable retroduplications in human populations suggests coupling of retrotransposition to cell division. *Genome Res*. **23**, 2042–2052 (2013).
36. Cooper, G. M. et al. A copy number variation morbidity map of developmental delay. *Nat. Genet*. **43**, 838–846 (2011).
37. Hook, E. B. & Hamerton, J. L. in *Population Cytogenetics: Studies in Humans* (eds Hook, E. B. & Porter, I. H.) 63–79 (Academic Press, 1977).
38. Forabosco, A., Percesepe, A. & Santucci, S. Incidence of non-age-dependent chromosomal abnormalities: a population-based study on 88965 amniocenteses. *Eur. J. Hum. Genet*. **17**, 897–903 (2009).
39. Malhotra, A. et al. Breakpoint profiling of 64 cancer genomes reveals numerous complex rearrangements spawned by homology-independent mechanisms. *Genome Res*. **23**, 762–776 (2013).
40. Conrad, D. F. et al. Mutation spectrum revealed by breakpoint sequencing of human germline CNVs. *Nat. Genet*. **42**, 385–391 (2010).
41. Quinlan, A. R. et al. Genome-wide mapping and assembly of structural variant breakpoints in the mouse genome. *Genome Res*. **20**, 623–635 (2010).
42. Mills, R. E. et al. Mapping copy number variation by population-scale genome sequencing. *Nature* **470**, 59–65 (2011).
43. Kidd, J. M. et al. A human genome structural variation sequencing resource reveals insights into mutational mechanisms. *Cell* **143**, 837–847 (2010).
44. Quinlan, A. R. & Hall, I. M. Characterizing complex structural variation in germline and somatic genomes. *Trends Genet*. **28**, 43–53 (2012).
45. Ruderfer, D. M. et al. Patterns of genic intolerance of rare copy number variation in 59,898 human exomes. *Nat. Genet*. **48**, 1107–1111 (2016).
46. Huang, N., Lee, I., Marcotte, E. M. & Hurles, M. E. Characterising and predicting haploinsufficiency in the human genome. *PLoS Genet*. **6**, e1001154 (2010).
47. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
48. Kundaje, A. et al. Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).
49. Khurana, E. et al. Integrative annotation of variants from 1092 humans: application to cancer genomics. *Science* **342**, 1235587 (2013).
50. Visel, A., Minovitsky, S., Dubchak, I. & Pennacchio, L. A. VISTA Enhancer Browser—a database of tissue-specific human enhancers. *Nucleic Acids Res*. **35**, D88–D92 (2007).
51. Fishilevich, S. et al. GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. *Database* **2017**, bax028 (2017).
52. Lesurf, R. et al. ORegAnno 3.0: a community-driven resource for curated regulatory annotation. *Nucleic Acids Res*. **44**, D126–D132 (2016).

NHGRI Centers for Common Disease Genomics

Goncalo R. Abecasis[15], Elizabeth Appelbaum[1], Julie Baker[16], Eric Banks[6], Raphael A. Bernier[17], Toby Bloom[10], Michael Boehnke[15], Eric Boerwinkle[9,18], Erwin P. Bottinger[19], Steven R. Brant[20], Esteban G. Burchard[21], Carlos D. Bustamante[16], Lei Chen[1], Judy H. Cho[19,22,23], Rajiv Chowdhury[24], Ryan Christ[1], Lisa Cook[1], Matthew Cordes[1], Laura Courtney[1], Michael J. Cutler[25], Mark J. Daly[6,26,27], Scott M. Damrauer[28], Robert B. Darnell[10,29,30], Tracie Deluca[1], Huyen Dinh[9], Harsha Doddapaneni[9], Evan E. Eichler[31,32], Patrick T. Ellinor[6,33], Andres M. Estrada[34], Yossi Farjoun[6], Adam Felsenfeld[35], Tatiana Foroud[36], Nelson B. Freimer[37], Catrina Fronick[1], Lucinda Fulton[1], Robert Fulton[1], Stacy Gabriel[6], Liron Ganel[1], Shailu Gargeya[10], Goren Germer[10], Daniel H. Geschwind[38,39,40], Richard A. Gibbs[9], David B. Goldstein[41,42], Megan L. Grove[9], Namrata Gupta[6], Christopher A. Haiman[43], Yi Han[9], Daniel Howrigan[6,27], Jianhong Hu[9], Carolyn Hutter[35], Ivan Iossifov[44], Bo Ji[1], Lynn B. Jorde[45], Goo Jun[18], John Kane[46], Chul Joo Kang[1], Hyun Min Kang[15], Sek Kathiresan[6,33,47], Eimear E. Kenny[19,22,48,49], Lily Khaira[10], Ziad Khan[9], Amit Khera[6,33,47], Charles Kooperberg[50], Olga Krasheninina[9], William E. Kraus[51], Subra Kugathasan[52], Markku Laakso[53], Tuuli Lappalainen[10,54], Adam E. Locke[1,3], Ruth J. F. Loos[19], Amy Ly[1], Robert Maier[6,27], Tom Maniatis[10,55], Loic Le Marchand[56], Gregory M. Marcus[57], Richard P. Mayeux[58], Dermot P. B. McGovern[59], Karla S. Mendoza[34], Vipin Menon[9], Ginger A. Metcalf[9], Zeineen Momin[9], Guiseppe Narzisi[10], Joanne Nelson[1], Caitlin Nessner[9], Rodney D. Newberry[3], Kari E. North[60], Aarno Palotie[6,26,27], Ulrike Peters[50], Jennifer Ponce[1], Clive Pullinger[46], Aaron Quinlan[45], Daniel J. Rader[61], Stephen S. Rich[62], Samuli Ripatti[6,26,27], Dan M. Roden[63], Veikko Salomaa[64], Jireh Santibanez[9], Svati H. Shah[51], M. Benjamin Shoemaker[63], Heidi Sofia[35], Taylorlyn Stephan[35], Christine Stevens[6], Stephan R. Targan[59], Marja-Riitta Taskinen[65,66], Kathleen Tibbetts[6], Charlotte Tolonen[6], Tychele Turner[31], Paul De Vries[18], Jason Waligorski[1], Kimberly Walker[9], Vivian Ota Wang[35], Michael Wigler[10,44], Richard K. Wilson[1,67], Lara Winterkorn[10], Genevieve Wojcik[16], Jinchuan Xing[12], Erica Young[1,3], Bing Yu[18] & Yeting Zhang[12]

[15]Department of Biostatistics and Center for Statistical Genetics, University of Michigan, School of Public Health, Ann Arbor, MI, USA. [16]Department of Genetics, Stanford University, Stanford, CA, USA. [17]Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, USA. [18]Human Genetics Center and Department of Epidemiology, University of Texas Health Science Center, Houston, TX, USA. [19]The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [20]Department of Medicine, Rutgers Robert Wood Johnson Medical School, Rutgers, The State University of New Jersey, New Brunswick, NJ, USA. [21]Department of Bioengineering, University of California, San Francisco, San Francisco, CA, USA. [22]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [23]Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [24]MRC/BHF Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK. [25]Intermountain Heart Institute, Intermountain Medical Center, Murray, UT, USA. [26]Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland. [27]Analytical and Translational Genetics Unit, Psychiatric and Neurodevelopmental Genetics Unit, Massachusetts General Hospital, Boston, MA, USA. [28]Department of Surgery, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [29]Laboratory of Molecular Neuro-Oncology, The Rockefeller University, New York, NY, USA. [30]Howard Hughes Medical Institute, The Rockefeller University, New York, NY, USA. [31]Department of Genome Science, University of Washington, Seattle, WA, USA. [32]Howard Hughes Medical Institute, University of Washington, Seattle, WA, USA. [33]Division of Cardiology, Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA. [34]National Laboratory of Genomics for Biodiversity (LANGEBIO), CINVESTAV, Irapuato, Mexico. [35]National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA. [36]Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA. [37]Center for Neurobehavioral Genetics, Jane and Terry Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, Los Angeles, CA, USA. [38]Program in Neurogenetics, Department of Neurology, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA. [39]Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA. [40]Institute of Precision Health, University of California, Los Angeles, Los Angeles, CA, USA. [41]Institute for Genomic Medicine, Columbia University Medical Center, New York, NY, USA. [42]Department of Genetics and Development, Columbia University Medical Center, New York, NY, USA. [43]Department of Preventative Medicine, University of Southern California, Los Angeles, CA, USA. [44]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA. [45]Department of Human Genetics, University of Utah, Salt Lake City, UT, USA. [46]Cardiovascular Research Institute, University of California, San Francisco, San Francisco, CA, USA. [47]Center for Genomic Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA. [48]The Icahn Institute of Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [49]Center for Statistical Genetics, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [50]Fred Hutchinson Cancer Research Center, Seattle, WA, USA. [51]Department of Medicine, Duke University, Durham, NC, USA. [52]Department of Pediatrics, Emory University School of Medicine, Atlanta, GA, USA. [53]Institute of Clinical Medicine, Internal Medicine, University of Eastern Finland, Kuopio, Finland. [54]Department of Systems Biology, Columbia University, New York, NY, USA. [55]Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY, USA. [56]Cancer Center, University of Hawaii, Honolulu, HI, USA. [57]Department of Medicine, University of California, San Francisco, San Francisco, CA, USA. [58]Department of Neurology, Columbia University, New York, NY, USA. [59]F. Widjaja Foundation Inflammatory Bowel and Immunobiology Research Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA. [60]Department of Epidemiology, University of North Carolina, Chapel Hill, NC, USA. [61]Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [62]Center for Public Health Genomics, University of Virginia School of Medicine, Charlottesville, VA, USA. [63]Department of Medicine, Vanderbilt University, Nashville, TN, USA. [64]National Institute for Health and Welfare, Helsinki, Finland. [65]Research Programs Unit, Diabetes and Obesity, University of Helsinki, Helsinki, Finland. [66]Heart and Lung Centre, Helsinki University Hospital, Helsinki, Finland. [67]Present address: Institute for Genomic Medicine, Nationwide Children's Hospital, Columbus, OH, USA. *A full list of authors and their affiliations appears in the Supplementary Information.

Methods

Generation of the B38 callset

Per-sample processing. This callset is derived from 23,559 individuals who were part of the CCDG programme as well as 950 Latino samples from the PAGE consortium. All data were produced at one of the four CCDG-funded sequencing centres and aligned to genome build GRCh38 using each individual centre's functionally equivalent pipeline implementation[25]. Per-sample calling was performed on 23,547 samples using LUMPY[23] (v.0.2.13), CNVnator[53] (v.0.3.3) and SVTyper[24] (v.0.1.4). We excluded human leukocyte antigen (HLA) sequences, decoy or alternate contigs and regions with copy number much higher than that expected (mean of 12 or more copies per genome across 409 samples) from SV calling with LUMPY (https://github.com/hall-lab/speedseq/blob/master/annotations/exclude.cnvnator_100bp.GRCh38.20170403.bed).

Per-sample quality control. We observed an excess of small (400–1,000-bp) singleton deletions (that is, present in only a single sample), suggesting a large number of false positives. On further investigation, this excess arose from differences between centres in library insert-size distribution. To reduce the number of false-positive small deletions, deletions of ≤1,000 bp were eliminated unless they had split read support in at least one sample. Subsequently, per-sample quality control was performed to eliminate outlier samples. We removed 213 samples in which variant counts (for any SV type) were >6 median absolute deviations from the median count for that type.

Merging and cohort-level re-genotyping. The remaining samples were processed into a single, joint callset using svtools[1] (https://github.com/hall-lab/svtools) (v.0.3.2), modified to allow for multi-stage merging. The code for this merging is available in a container hosted on DockerHub (https://hub.docker.com/r/ernfrid/svtools_merge_beta) (ernfrid/svtools_merge_beta:292bd3). Samples were merged using svtools lsort followed by svtools lmerge in batches of 1,000 samples (or fewer) within each cohort. The resulting per-cohort batches were then merged again using svtools lsort and svtools lmerge to create a single set of variants for the entire set of 23,331 remaining samples. This site list was then used to genotype each candidate site in each sample across the entire cohort using SVTyper (v.0.1.4). Genotypes for all samples were annotated with copy-number information from CNVnator. Subsequently, the per-sample VCFs were combined together using svtools vcfpaste. The resulting VCF was annotated with allele frequencies using svtools afreq, duplicate SVs were pruned using svtools prune, variants were reclassified using svtools classify (large sample mode) and any identical lines were removed. For reclassification of chromosomes X and Y, we used a container hosted on DockerHub (https://hub.docker.com/r/ernfrid/svtools_classifier_fix) (ernfrid/svtools_classifier_fix:v1). All other steps to assemble the cohort above used the same container that was used for merging.

Callset tuning. Using the variant calling control trios, we chose a mean sample quality (MSQ) cut-off for INV and BND variant calls that yielded a Mendelian error rate of approximately 5%. INVs passed if: MSQ ≥ 150; neither split-read nor paired-end LUMPY evidence made up more than 10% of total evidence; each strand provided at least 10% of read support. BNDs passed if MSQ ≥ 250.

Genotype refinement. MEI and DEL genotypes were set to missing on a per-sample basis (https://github.com/hall-lab/svtools/blob/develop/scripts/filter_del.py, commit 5c32862) if the site was poorly captured by split reads. Genotypes were set to missing if the size of the DEL or MEI was smaller than the minimum size discriminated at 95% confidence by SVTyper (https://github.com/hall-lab/svtools/blob/develop/scripts/del_pe_resolution.py, commit 3fc7275). DEL and MEI genotypes for sites with allele frequency ≥ 0.01 were refined based on clustering of allele balance and copy-number values within the datasets produced by each sequencing centre (https://github.com/hall-lab/svtools/blob/develop/scripts/geno_refine_12.py, commit 41fdd60). In addition, duplications were re-genotyped with more-sensitive parameters to better reflect the expected allele balance for simple tandem duplications (https://github.com/ernfrid/regenotype/blob/master/resvtyper.py, commit 4fadcc4).

Filtering for size. The remaining variants were filtered to meet the size definition of a SV (≥50 bp). The length of intra-chromosomal generic BNDs was calculated using vawk (https://github.com/cc2qe/vawk) as the difference between the reported positions of each breakpoint.

Large callset sample quality control. Of the remaining samples, we evaluated per-sample counts of deletions, duplications and generic BNDs within the low-allele-frequency (0.1%–1%) class. Samples with variant counts exceeding 10 median absolute deviations from the mean for any of the 3 separate variant classes were removed. In addition, we removed samples with genotype missingness >2%. These quality control filters removed a total of 120 additional samples. Finally, we removed 64 samples that were identified as duplicates or twins in a larger set of data.

Breakpoint resolution

Breakpoint resolution was calculated using BCFtools (v.1.3.1) query to create a table of confidence intervals for each variant in the callset, but excluding secondary BNDs. Each breakpoint contains two 95% confidence intervals, one each around the start location and end location. Summary statistics were calculated in RStudio (v.1.0.143; R v.3.3.3).

Self-reported ethnicity

Self-reported ethnicity was provided for each sample via the sequencing centre and aggregated by the NHGRI Genome Sequencing Program (GSP) coordinating centre. For each combination of reported ethnicity and ancestry, we assigned a super-population, continent (based on the cohort) and ethnicity. Samples in which ancestry was unknown, but the sample was Hispanic, were assigned to the Americas (AMR) super-population. Summarized data are presented in Extended Data Table 1.

Sample relatedness

As SNV calls were not yet available for all samples at the time of the analysis, relatedness was estimated using large (>1 kb), high-quality autosomal deletions and MEIs with allele frequency >1%. These were converted to plink format using PLINK (v.1.90b3.38) and then subjected to kinship calculation using KING[54] (v.2.0). The resulting output was parsed to build groups of samples connected through first-degree relationships (kinship coefficient > 0.177). Correctness was verified by the successful recapitulation of the 36 complete Coriell trios included as variant calling controls. We note that, in analyses of the full B38 callset (which contains cohorts of families), 'ultra-rare' or 'singleton' variants were defined as those unique to a family. For analyses of the of 4,298 sample subset of unrelated individuals with both SV and SNP/indel calls, 'singleton' variants were defined as those present as a single allele.

Callset summary metrics

Callset summary metrics were calculated by parsing the VCF files with BCFtools (v.1.3.1) query to create tables containing information for each variant–sample pairing or variant alone, depending on the metric. Breakdowns of the BND class of variation were performed using vawk to calculate orientation classes and sizes. These were summarized using Perl and then transformed and plotted using RStudio (v.1.0.143; R v.3.3.3).

Ultra-rare variant analysis

We defined an ultra-rare variant as any variant unique to one individual or one family of first-degree relatives. We expect the false-positive rate of ultra-rare variants to be low because systematic false positives owing to alignment issues are likely to be observed in multiple unrelated individuals. Therefore, we considered both high- and low-confidence variants in all ultra-rare analyses.

Constructing variant chains. Complex variants were identified as described previously[4] by converting each ultra-rare SV to bed format and, within a given family, clustering breakpoints occurring within 100,000 bp of each other using BEDTools[55] (v.2.23.0) cluster. Any clusters linked together by BND variants were merged together. The subsequent collection of variant clusters and linked variant clusters (hereafter referred to as chains) were used for both retrogene and complex variant analyses.

Manual review. Manual review of variants was performed using the Integrative Genomics Viewer (IGV) (v.2.4.0). Variants were converted to BED12 using svtools (v.0.3.2) for display within IGV. For each sample, we generated copy-number profiles using CNVnator (v.0.3.3) in 100-bp windows across all regions contained in the variant chains.

Retrogene insertions. Retrogene insertions were identified by examining the ultra-rare variant chains constructed as described above. For each chain, we identified any constituent SV with a reciprocal overlap of 90% to an intron using BEDTools (v.2.23.0). For each variant chain, the chain was deemed a retrogene insertion if it contained one or more BND variants with +/− strand orientation that overlapped an intron. In addition, we flagged any chains that contained non-BND SV calls, as their presence was indicative of a potential misclassification, and manually inspected them to determine whether they represented a true retrogene insertion.

Complex variants. We retained any cluster(s) incorporating three or more SV breakpoint calls, but removed SVs identified as retrogene insertions either during manual review or algorithmically. In addition, we excluded one call deemed to be a large, simple variant after manual review.

Large variants. Ultra-rare variants >1 Mb in length were selected and any overlap with identified complex variants identified and manually reviewed. Of five potential complex variants, one was judged to be a simple variant and included as a simple variant, whereas the rest were clearly complex variants and excluded. Gene overlap was determined as an overlap ≥ 1 bp with any exon occurring within protein-coding transcripts from Gencode v.27 marked as a principal isoform according to APPRIS[56].

Balanced translocations. Ultra-rare generic BND variants, of any confidence class, connecting two chromosomes and with support (>10%) from both strand orientations were initially considered as candidate translocations. We further filtered these candidates to require exactly two reported strand orientations indicating reciprocal breakpoints (that is, +−/−+, −+/+−, −−/++, ++/−−), no read support from any sample with a homozygous reference genotype, at least one split read supporting the translocation from samples containing the variant, and <25% overlap of either breakpoint with any simple repeat (downloaded from ftp://hgdownload.cse.ucsc.edu/goldenPath/hg38/database/simpleRepeat.txt.gz).

Comprehensive annotations from the Gencode v.27 GTF (ftp://ftp.ebi.ac.uk/pub/databases/gencode/Gencode_human/release_27/gencode.v27.annotation.gtf.gz) were used to determine the number of affected genes. A BED file of all introns was created by converting transcripts and exons to BED entries and subtracting all exons from their respective transcripts using BEDTools (v.2.23.0). To identify translocations affecting genes, the translocations were converted to BEDPE using svtools (v.0.3.1), padded by 1 bp and intersected with introns using BEDTools (v.2.23.0). The number of unique chromosome–gene name pairs for each translocation was used to determine the number of affected genes affected by each breakpoint.

To determine whether a translocation resulted in an in-frame fusion, we converted to BEDPE, padded by 1 bp and intersected the breakpoints with all introns using BEDTools (v.2.23.0). Each intron entry was then padded by 1 bp and intersected with the Gencode GTF file using BEDTools (v.2.23.0) and restricting to coding exons of the same transcript as the intron. Then, for each set of exons intersected by a given translocation, all combinations of transcripts were compared, taking into account their orientation and the orientation of the breakpoint, to determine whether the frame was maintained across the potentially fused exons. The resulting two candidate translocations were manually reviewed by reconstructing the transcript sequence of the fusion and translating the resulting DNA sequence using https://web.expasy.org/translate/ to confirm a single open-reading frame was maintained.

Generation of the B37 callset

Per-sample processing. This callset was constructed starting from a set of 8,455 individuals: 8,181 samples from 8 cohorts sequenced at the McDonnell Genome Institute, as well as 274 samples from the Simons Genome Diversity Project downloaded from EMBL-EBI (https://www.ebi.ac.uk/ena/data/view/PRJEB9586). All samples passed standard production quality control metrics and had a mean depth of coverage >20×. Data were aligned to GRCh37 using the SpeedSeq (v0.1.2) realignment pipeline. Per-sample SV calling was performed with SpeedSeq sv (v.0.1.2) using LUMPY (v.0.2.11), CNVnator-multi and SVTyper (v.0.1.4) on our local compute cluster. For LUMPY SV calling, we excluded high-copy-number outlier regions derived from >3,000 Finnish samples as described previously[1] (https://github.com/hall-lab/speedseq/blob/master/annotations/exclude.cnvnator_100bp.112015.bed).

Per-sample quality control. Following a summary of per-sample counts, samples with counts of any variant class (DEL, DUP, INV or BND) exceeding the median plus 10 times the median absolute deviation for that class were excluded from further analysis; 17 such samples were removed.

Merging. The remaining samples were processed into a single, joint callset using svtools (v.0.3.2) and the two-stage merging workflow (as described above): each of the nine cohorts was sorted and merged separately in the first stage, and the merged calls from each cohort sorted and merged together in the second stage.

Cohort-level re-genotyping. The resulting SV loci were then re-genotyped with SVTyper (v.0.1.4) and copy-number annotated using svtools (v.0.3.2) in parallel, followed by a combination of single-sample VCFs, frequency annotation and pruning using the standard workflow for svtools (v.0.3.2). A second round of re-genotyping with more-sensitive parameters to better reflect the expected allele balance for simple tandem duplications (https://github.com/ernfrid/regenotype/blob/master/resvtyper.py, commit 4fadcc4) was then performed, followed by another round of frequency annotation, pruning and finally reclassification using svtools (v.0.3.2) and the standard workflow.

Callset tuning and site-level filtering. Genotype calls for samples in 452 self-reported trios were extracted, and Mendelian error rates calculated using a custom R script; we counted as a Mendelian error any child genotype inconsistent with inheritance of exactly one allele from the mother and exactly one allele from the father. Filtering was performed as described for the B38 callset: INVs passed if: MSQ ≥ 150;

neither split-read nor paired-end LUMPY evidence made up <10% of total evidence; each strand provided at least >10% of read support. Generic BNDs passed if MSQ ≥ 250. SVs of length <50 bp were removed, according to our working definition of 'structural variation'.

Final sample-level filtering. Nine samples with retracted consents, and two hydatidiform mole samples were removed from the callset. Subsequently, the numbers of quality-control-passing, very rare (<0.1% MAF) DELs, DUPs and BNDs per sample were determined. Excluding the samples in the Simons Genome Diversity cohort (which were expected, in general, to have unusually high counts of rare variants), we determined the median and median absolute deviation (MAD) of the per-sample counts of each type, and excluded outlier samples with a count exceeding the median + 10 × MAD of any type. Nine samples were removed in this way. Finally, kinship was estimated using KING (v.2.0) based on high-quality, autosomal deletion and MEI calls with population allele frequency >1%. Each SV was annotated in the VCF according to the number of distinct, first-degree family clusters in which it was observed, as for the B38 callset.

PCA. A set of unrelated individuals (containing no first- or second-degree relatives) was extracted using KING (v.2.0). PCA was performed using smartpca (v.13050) on a VCF of all high-quality DEL and MEI variant calls with population allele frequency >1%. Eigenvectors were estimated based on the set of unrelated samples, and then all samples projected onto the eigenvectors.

Generation of the B38 SNV and indel callset and quality control

Per-sample calling was performed at the Broad Institute as part of CCDG joint-calling of 22,609 samples using GATK[57,58]. HaplotypeCaller v.3.5-0-g36282e4. All samples were joint-called at the Broad Institute using GATK v.4.beta.6, filtered for sites with an excess heterozygosity value of more than 54.69 and recalibrated using VariantRecalibrator with the following features: QD, MQRankSum, ReadPosRankSum, FS, MQ, SOR and DP. Individual cohorts were subset out of the whole CCDG callset using Hail v.0.2 (https://github.com/hail-is/hail). After SNV and indel variant recalibration, multi-allelic variants were decomposed and normalized with vt (v.0.5)[59]. Duplicate variants and variants with symbolic alleles were then removed. Afterwards, variants were annotated with custom computed allele balance statistics, 1000 Genomes Project allele frequencies[28], gnomAD-based population data[27], VEP (v.88)[60], CADD[29] (v.1.2) and LINSIGHT[30]. Variants having greater than 2% missingness were soft-filtered. Samples with high rates of missingness (>2%) or with mismatches between reported and genetically estimated sex (determined using PLINK v.1.90b3.45 sex-check) were excluded. The LOFTEE plug-in (v.0.2.2-beta; https://github.com/konradjk/loftee) was used to classify putative loss-of-function SNVs and indels as high or low confidence.

Annotation of gene-altering SV calls

The VCF was converted to BEDPE format using svtools vcftobedpe The resulting BEDPE file was intersected (using BEDTools (v.2.23.0) intersect and pairtobed) with a BED file of coding exons from Gencode v.27 with principal transcripts marked according to APPRIS[56]. The following classes of SV were considered as putative gene-altering events: (1) DEL, DUP, or MEI intersecting any coding exon; (2) INV intersecting any coding exon and with either breakpoint located within the gene body; and (3) BND with either breakpoint occurring within a coding exon.

Gene-based estimation of dosage sensitivity

We followed a previously described method[45], to estimate genic dosage sensitivity scores using counts of exon-altering deletions and duplications in a combined callset comprising the 14,623 sample pan-CCDG callset plus 3,172 non-redundant samples from the B37 callset. B37 CNV calls were lifted over to B38 as BED intervals using CrossMap (v.0.2.1)[61]. We determined the counts of deletions and duplications that intersect coding exons of principal transcripts of any autosomal gene. In the previous study[45], the expected number of CNVs per gene was modelled as a function of several genomic features (GC content, mean read depth and so on), some of which were relevant to their exome read-depth CNV callset but not to our WGS-based breakpoint mapping lumpy/svtools callset. To select the relevant features for prediction, using the same set of gene-level annotations as described previously[45], we restricted to the set of genes in which fewer than 1% of samples carried an exon-altering CNV, and used l^1-regularized logistic regression (from the R glmnet package[62], v.2.0-13), with the penalty λ chosen by tenfold cross-validation. The selected parameters (gene length, number of targets and segmental duplications) were then used as covariates in a logistic regression-based calculation of per-gene intolerance to DEL and DUP, similar to that described previously[45]. For deletions (or duplications, respectively), we restricted to the set of genes with <1% of samples carrying a DEL, to estimate the parameters of the logistic model. We then applied the fitted model to the full set of genes to calculate genic CNV intolerance scores as the residuals of the logistic regression of CNV frequency on the genomic features, standardized as z-scores and with winsorization of the lower 5th percentile.

Genome-wide estimation of deleterious variants

To estimate the relative numbers of deleterious SNVs, indels, DELs and DUPs genome-wide in the normal population, we relied on a subset of 4,298 samples from the B38 callset for which we had joint variant callsets for both SNVs/indels (GATK) and SVs (lumpy/svtools). Each SNV and indel was annotated with CADD[29] and LINSIGHT[30] scores as described above. CADD and LINSIGHT scores were converted to percentiles and singleton rates (where 'singleton' was defined as a variant present as a single allele) calculated for variants above each score threshold. CADD and LINSIGHT scores were then calibrated to a standard scale by matching singleton rates. Each DEL and DUP was annotated with CADD and LINSIGHT scores, calculated as the mean of the top 10 single-base CADD or LINSIGHT scores, respectively, for the span of the CNV (similar to SVScore[63]). The CNV-level CADD and LINSIGHT scores were then standardized using the above calibration curves. Finally, each variant (SNV, indel or CNV) was assigned a combined CADD–LINSIGHT score, calculated as the maximum of the two distinct scores.

The combined scores provided a means to rank, within each variant class, variants in order of deleteriousness. We calculated the singleton rate for the set of all LOFTEE high confidence protein-truncating SNVs and indels in autosomal genes. We then estimated the number of deleterious variants of each type genome-wide by choosing the combined CADD–LINSIGHT score threshold as the minimum value, such that the singleton rate for the set of higher-scoring variants was greater than or equal to the singleton-rate for LOFTEE high-confidence protein-truncating variants.

Annotation of noncoding elements

We divided the genome into 1-kb non-overlapping windows to investigate the rates of CNV occurrence relative to various classes of coding and noncoding elements, genome-wide. Windows intersecting assembly gaps or high-copy-number outlier regions (as described above) and windows with fewer than 50% of bases uniquely mappable as determined using GEM-mappability (build 1.315)[64] were excluded from analysis. BED tracks of genomic annotations for the noncoding dosage sensitivity analysis were created as described below.

The phastcons-20way[65] conservation track was downloaded from the UCSC genome browser (rsync://hgdownload.cse.ucsc.edu/goldenPath/hg38/phastCons20way/hg38.phastCons20way.wigFix.gz) and converted to bed format. The mean PhastCons score for each 1-kb window was calculated using BEDTools map. Quantiles of mean window-level PhastCons scores were calculated and used as thresholds for the sensitivity analysis.

The LINSIGHT[30] score track was downloaded from CSHL (http://compgen.cshl.edu/LINSIGHT/LINSIGHT.bw). The 1-kb genomic windows were lifted over to hg19 using CrossMap (v.0.2.1), annotated with mean per-window LINSIGHT scores using BEDTools map and lifted back to GRChb38. Quantiles of mean window-level LINSIGHT scores were calculated and used as thresholds for the sensitivity analysis.

Genehancer[51] enhancers were downloaded from GeneCards (https://genecards.weizmann.ac.il/geneloc/index.shtml) and converted to bed format.

Vista[50] enhancers were downloaded from LBL (https://enhancer.lbl.gov/cgi-bin/imagedb3.pl?page_size=20000;show=1;search.result=yes;page=1;form=search;search.form=no;action=search;search.sequence=1), restricted to human enhancers, converted to bed format and lifted over to GRChb38 using CrossMap.

Encode[47] DNase hypersensitivity sites and transcription-factor-binding sites were downloaded from UCSC (http://hgdownload.cse.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeRegDnaseClustered/wgEncodeRegDnaseClusteredV3.bed.gz, http://hgdownload.cse.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeRegTfbsClustered/wgEncodeRegTfbsClusteredV3.bed.gz) and lifted over to GRCHb38 using CrossMap.

Oreganno[66] literature-curated enhancers were downloaded from UCSC (http://hgdownload.cse.ucsc.edu/goldenpath/hg19/database/oreganno.txt.gz) converted to bed format and lifted over to GRChb38 using CrossMap.

Sensitive[49], transcription-factor-bound, ultra-conserved[67] and HOT[68] regions were downloaded from the funseq2[69] resources (http://archive.gersteinlab.org/funseq2.1.0_data).

Dragon enhancers were downloaded from DENdb[70] (http://www.cbrc.kaust.edu.sa/dendb/src/enhancers.csv.zip), converted to bed format, lifted over to GRChb37 and filtered for score >2.

Chromatin interaction domains derived from Hi-C on human ES cell and IMR90 cells[71] were downloaded from http://compbio.med.harvard.edu/modencode/webpage/hic/, and distances between adjacent topological domains were calculated with BEDTools. When the physical distance between adjacent topological domains was <400 kb, these were classified as TAD boundaries; otherwise, they were classified as unorganized chromatin. The TAD boundaries and unorganized chromatin data were converted to bed format and lifted over to GRCh38 using CrossMap.

Roadmap chromatin state segmentations for 127 epigenomes were downloaded from Roadmap[48] (https://egg2.wustl.edu/roadmap/data/byFileType/chromhmmSegmentations/ChmmModels/coreMarks/jointModel/final/) and lifted over to GRCh38. BEDTools multiinter was used to determine the number of epigenomes in which each segment was present.

Dosage sensitivity of noncoding elements

To maximize power, DEL and DUP calls from the non-redundant combination of the B37 and B38 callsets (as described above) were used for this analysis. Each window was further characterized by its distance to the nearest exon (the minimum distance between any point in the window and any point in the exon) and the pLI score of the gene corresponding to the nearest exon. The pLI score was set to zero for genes with pLI undefined. In the event that exons of two genes were equidistant to the window, the max of the two pLI scores was selected.

For a given SV type (DUP or DEL) and a given functional annotation (for example, VISTA enhancers), each window was characterized by the presence or absence of one or more SV and the presence or absence of one of more genomic features. We observed a depletion of CNVs in windows near exons, and in particular near exons of loss-of-function-intolerant genes (Fig. 4a). As such, we used a Cochran–Mantel–Haenszel estimate of the odds ratios for each SV type or functional annotation, while stratifying for the proximity to the nearest exon as well as that exon's loss-of-function intolerance score (pLI). Because adjacent windows are not strictly independent observations—that is, CNV or features may overlap adjacent windows, inducing some spatial correlations—we used a block bootstrap method (resampling was performed on blocks of 10 windows) to estimate robust confidence intervals.

Long-read validation

PacBio long-read sequences from nine 1000 Genomes Project (1KG) samples sequenced to deep coverage (>68-87x) at the McDonnell Genome Institute were used as an orthogonal means of validating SV calls.These PacBio data are available in SRA (see accessions in Supplementary Table 2) and were generated independently from the long-read data used by the Human Genome Structural Variation Consortium (HGSVC) to create the long-read SV callset used for sensitivity analyses described below[26]. The long-read sequences were aligned to GRCh38 using minimap2 (ref. [72]) (v.2.16-r922; parameters -ax map-pb). Split-read alignments indicating putative SVs were converted to BEDPE format[55] as described previously[23,24,73]. Similarly, deletions or insertions longer than 50 bp contained within PacBio reads (as determined based on the cigar strings) were converted to BEDPE format. We used BEDTools to judge the overlap between short-read SV calls and the long-read alignments. We judged an SV call to be validated when ≥ 2 long-reads exhibited split-read mappings in support of the SV call. For a long-read mapping to support an SV call, we required that it must predict a consistent SV type (for example, deletion) and exhibit substantial physical overlap with the SV call, where overlap can be met by either of the following criteria: (1) the two breakpoint intervals predicted by the SV call and the two breakpoint intervals predicted by the long-read split-read mapping overlap with each other on both sides, as determined by BEDTools pairtopair using 100 bp of "slop" (-type -is both -slop 100); or (2) the SV call and the long-read split-read (or cigar-derived indel variant) exhibit 90% reciprocal overlap with one another (BEDTools intersect -r -f 0.9). The above criteria for SV validation based on long-read support were selected based on extensive manual review of SV calls in the context of supporting data including read-depth profiles and long-read mappings from all nine samples, and are the basis for the validation rates reported in the main text and in Supplementary Table 3. However, we also show the range of validation rates that are obtained when using more lenient or strict measures of physical overlap, and when requiring a varying number of supporting PacBio reads (Extended Data Fig. 5), in both carriers and non-carriers of SVs from various classes. We also note that 3 of the 6 singleton SV calls that are not validated by long reads appear to be true variants based on manual review of read-level evidence, in which it appears that long-reads failed to validate true short-read SV calls owing to subtle differences in how coordinates were reported at local repeats. Our false discovery rate estimates may be conservative owing to these effects.

To conduct a comparison to HGSVC using the three samples shared between our datasets (NA19240, HG00514, HG00733), all non-reference, autosomal SV calls for each of the three samples were extracted from the CCDG B38 and HGSVC[26] Illumina short-read callsets. For HGSVC variants detected solely by read-depth analysis, for which genotype information was not available, a variant was defined to be non-reference if its predicted copy-number differed from the mode for that site across the nine samples in that callset (which includes the parents of NA19240, HG00514 and HG00733). The short-read calls from our study and HGSVC for the three relevant samples were converted to BEDPE format using svtools vcftobedpe. The three single-sample VCFs from the HGSVC PacBio long-read SV callset were converted to BEDPE format in similar fashion. For HGSVC Illumina calls (which had been taken from a callset comprising three trios, rather than a large cohort) variants were classified as rare if seen in only one of the six trio founders and either absent from or observed at frequency <1% in the 1KG phase 3 SV callset.

Long-read SV truth set construction

To evaluate the sensitivity of our callset, we constructed a high-confidence truth set from the comprehensive HGSVC long-read SV callset created using reference-guided de novo assembly[26]. The assembly-based long-read truth set includes all autosomal SVs reported by HGSVC[26] that were also validated by split-read alignments from the PacBio data generated independently at our centre. Here, an HGSVC call was judged to be validated by long-read data when two or more long reads exhibited split-read mappings or cigar-derived SV calls that match the HGSVC call in terms of the predicted SV type and breakpoint intervals, allowing 100 bp of "slop" to account for positional uncertainty (BEDTools pairtopair -type -is both -slop 100). To account for the variant classification scheme of the HGSVC callset—which only has two variant categories, INS and DEL—we allowed INS variants to be validated by long reads suggesting either insertion or tandem duplication variants. Variants were classified as STRs if either >50% of sequence from both reported breakpoint intervals or >50% of sequence contained in the outer span of the variant overlapped a GRCh38 track of simple repeats downloaded from the UCSC Table Browser. The interval spanned by each variant was converted to bed format and lifted over to hg19 using CrossMap. A combined CADD–LINSIGHT score was calculated for each variant based on the mean of the top 10 CADD-scoring and the mean of the top 10 LINSIGHT-scoring positions, as described in the section 'Genome-wide estimation of deleterious variants'.

Lifting over of the 1KG phase3 SV callset

The 1KG phase 3 SV callset was lifted over from GRCh37 to GRCh38 by first converting to BEDPE format using svtools vcftobedpe. The outer span of each variant was then converted to bed format and lifted over using CrossMap[61]. For SVs that were not lifted over as contiguous intervals, discontiguous regions within 1 kb were merged using BEDTools merge, and the largest of the merged variants were selected. The lifted-over bed interval was then converted back to BEDPE by padding each endpoint with 100 bp.

Assessment of sensitivity using the HGSVC long-read truth set

Sensitivity of the CCDG B38 and HGSVC Illumina short-read callsets to detect variants in the HGSVC long-read truth set was determined by converting each single-sample VCF to BEDPE format using svtools vcftobedpe and calculating overlaps using BEDTools pairtopair, allowing for 100 bp of "slop". For DEL calls, a variant was considered to be detected only if both breakpoints overlapped, and the type of the overlapping call was consistent with a deletion (that is, DEL, MEI, CNV or BND). For INS calls in the long-read callset, variants were considered to be detected if either breakpoint overlapped and the overlapping call was consistent with an insertion (that is, DUP, INS, CNV, MEI or BND).

Comparison with the 1KG phase 3 SV callset[5] necessitated the use of a slightly different sensitivity metric, as 1KG analysed the parents of HG00733 and NA19249, but not the trio offspring themselves. As, with rare exception, germline variants present in the child should also be present in one of the two parents, the rate at which HGSVC long-read calls in the truth set were detected in at least one parent in each of the CCDG B38, HGSVC and 1KG callsets serves as an informative alternate measure of 'sensitivity'.

Genotype comparison to 1KG

Genotype comparisons were performed for the five parental samples (NA19238, NA19239, HG00513, HG00731 and HG00732) present in both the CCDG B38 and the 1KG phase 3 SV callsets. Each callset was subset (using BCFtools) to the set of autosomal SVs with a non-reference call in at least one of the five parental samples and converted to BEDPE format. Variants in the 1KG callset detected using read-depth methods only were excluded. Bedtools pairtopair (100 bp slop, overlap at both breakpoints) was used to determine the set of variants called in both the five-sample CCDG callset and the five-sample 1KG callset, requiring consistent SV type. For each variant site in each sample, genotypes from the two callsets were compared. Results were tallied, and concordance rates and kappa statistics ('irr' package) were calculated in R.

Pedigree analysis

Pedigree analyses were performed on three-generation pedigrees from Utah collected as part of the Centre d'Étude du Polymorphisme Humain (CEPH) consortium. The analyses used a set of 576 CEPH samples contained in the B37 callset that remained after excluding 21 samples that had been deemed low-quality and/or possibly contaminated based on analysis of a SNV–indel callset (data not shown). The remaining samples comprise 409 trios, which were used in the estimation of transmission rates. The counts of all high-quality SVs called heterozygous in one parent, homozygous reference in the other and non-missing in the offspring were used to estimate transmission rates by frequency class, with Wilson score confidence intervals calculated using R binconf.

Mendelian errors for all high-quality (filter = PASS) SVs were calculated using PLINK (v.1.90b3.45), with the output restricted to variant-trio observations in which all three genotypes (father, mother and offspring) were non-missing. For each sample in the third generation (the 'F_2'; see Extended Data Fig. 6a) of any of the CEPH kindreds, Mendelian errors were counted by frequency class. The Mendelian error rate was calculated as the total number of Mendelian errors divided by the total number of non-reference, non-missing genotypes in F_2 generation samples for variants of that frequency class. De novo variants were defined as variants private to a single family in which both parental genotypes are 0/0 and the offspring genotype either 0/1 or 1/1, and were obtained by parsing the PLINK output. (Note that these variant counts are used as callset quality metrics and do not necessarily represent true de novo mutations.)

Transmission rates for putative de novo variants were calculated by restricting to all high-quality autosomal variants heterozygous in a second generation ('F_1') sample and homozygous reference in both of his/her parents ('P_0' generation) and his/her F_1 spouse. Each such variant was classified as transmitted if carried by any F_2 offspring, with transmission rates calculated as the number of transmitted variants out of the total. 'Missed heterozygous calls' were counted as the set of all family-private variants non-reference in at least two F_2 offspring siblings, but homozygous reference in both of the F_1 parents. The rate of missed heterozygous calls was calculated by dividing this count by the total count of family-private variants carried by at least two F_2 offspring siblings.

Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The sequencing data can be accessed through dbGaP (https://www.ncbi.nlm.nih.gov/gap) under the accession numbers provided in Supplementary Table 7. PacBio long-read data used for SV validation can be accessed through the Sequence Read Archive (SRA), under the accession numbers provided in Supplementary Table 2. The set of high-confidence HGSVC long-read-derived SV calls, validated by our independent PacBio data and used as a truth set, can be found in Supplementary File 3. Supplementary Files 1–4 can be found at https://github.com/hall-lab/sv_paper_042020.

Code availability

Custom code used in the long-read validation can be found here: https://github.com/abelhj/long-read-validation/tree/master.

53. Abyzov, A., Urban, A. E., Snyder, M. & Gerstein, M. CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res*. **21**, 974–984 (2011).
54. Manichaikul, A. et al. Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867–2873 (2010).
55. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
56. Rodriguez, J. M. et al. APPRIS: annotation of principal and alternative splice isoforms. *Nucleic Acids Res*. **41**, D110–D117 (2013).
57. DePristo, M. A. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet*. **43**, 491–498 (2011).
58. Poplin, R. et al. Scaling accurate genetic variant discovery to tens of thousands of samples. Preprint at bioRxiv https://www.biorxiv.org/content/10.1101/201178v3 (2018).
59. Tan, A., Abecasis, G. R. & Kang, H. M. Unified representation of genetic variants. *Bioinformatics* **31**, 2202–2204 (2015).
60. McLaren, W. et al. The Ensembl Variant Effect Predictor. *Genome Biol*. **17**, 122 (2016).
61. Zhao, H. et al. CrossMap: a versatile tool for coordinate conversion between genome assemblies. *Bioinformatics* **30**, 1006–1007 (2014).
62. Friedman, J., Hastie, T. & Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw*. **33**, 1–22 (2010).
63. Ganel, L., Abel, H. J. & Hall, I. M. SVScore: an impact prediction tool for structural variation. *Bioinformatics* **33**, 1083–1085 (2017).
64. Derrien, T. et al. Fast computation and applications of genome mappability. *PLoS One* **7**, e30377 (2012).
65. Siepel, A. et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res*. **15**, 1034–1050 (2005).
66. Griffith, O. L. et al. ORegAnno: an open-access community-driven resource for regulatory annotation. *Nucleic Acids Res*. **36**, D107–D113 (2008).
67. Bejerano, G. et al. Ultraconserved elements in the human genome. *Science* **304**, 1321–1325 (2004).
68. Yip, K. Y. et al. Classification of human genomic regions based on experimentally determined binding sites of more than 100 transcription-related factors. *Genome Biol*. **13**, R48 (2012).
69. Fu, Y. et al. FunSeq2: a framework for prioritizing noncoding regulatory variants in cancer. *Genome Biol*. **15**, 480 (2014).
70. Ashoor, H., Kleftogiannis, D., Radovanovic, A. & Bajic, V. B. DENdb: database of integrated human enhancers. *Database)* **2015**, bav085 (2015).
71. Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
72. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
73. Faust, G. G. & Hall, I. M. YAHA: fast and flexible long-read alignment with optimal breakpoint detection. *Bioinformatics* **28**, 2417–2424 (2012).

Acknowledgements We thank staff at the NHGRI for supporting this effort. This study was funded by NHGRI CCDG awards to Washington University in St Louis (UM1 HG008853), the Broad Institute of MIT and Harvard (UM1 HG008895), Baylor College of Medicine (UM1 HG008898) and New York Genome Center (UM1 HG008901); an NHGRI GSP Coordinating Center grant to Rutgers (U24 HG008956); and a Burroughs Wellcome Fund Career Award to I.M.H. Additional data production at Washington University in St Louis was funded by a separate NHGRI award (5U54HG003079). We thank S. Sunyaev for comments on the manuscript; T. Teshiba for coordinating samples for FINRISK and EUFAM sequencing; and the staff and participants of the ARIC study for their contributions; and we acknowledge all individuals who were involved in the collection of samples that were analysed for this study.

Data production for EUFAM was funded by 4R01HL113315-05; the Metabolic Syndrome in Men (METSIM) study was supported by grants to M. Laakso from the Academy of Finland (no. 321428), the Sigrid Juselius Foundation, the Finnish Foundation for Cardiovascular Research, Kuopio University Hospital and the Centre of Excellence of Cardiovascular and Metabolic Diseases supported by the Academy of Finland; data collection for the CEPH pedigrees was funded by the George S. and Dolores Doré Eccles Foundation and NIH grants GM118335 and GM059290; study recruitment at Washington University in St Louis was funded by the DDRCC (NIDDK P30 DK052574) and the Helmsley Charitable Trust; study recruitment at Cedars-Sinai was supported by the F. Widjaja Foundation Inflammatory Bowel and Immunobiology Research Institute, NIH/NIDDK grants P01 DK046763 and U01 DK062413 and the Helmsley Charitable Trust; study recruitment at Intermountain Medical Center was funded by the Dell Loy Hansen Heart Foundation; the Late Onset Alzheimer's Disease Study (LOAD) study was funded by grants to T. Foroud (U24AG021886, U24AG056270, U24AG026395 and R01AG041797); the Atherosclerosis Risk in Communities (ARIC) study was funded by the NHLBI (HHSN268201700001I, HHSN268201700002I, HHSN268201700003I, HHSN268201700004I and HHSN268201700005I); and the PAGE programme is funded by the NHGRI with co-funding from the NIMHD (U01HG007416, U01HG007417, U01HG007397, U01HG007376 and U01HG007419). Samples from the BioMe Biobank were provided by The Charles Bronfman Institute for Personalized Medicine at the Icahn School of Medicine at Mount Sinai. The Hispanic Community Health Study/Study of Latinos was carried out as a collaborative study supported by the NHLBI (N01-HC65233, N01-HC65234, N01-HC65235, N01-HC65236 and N01-HC65237), with contributions from the NIMHD, NIDCD, NIDCR, NIDDK, NINDS and NIH ODS. The Multiethnic Cohort (MEC) study is funded through the NCI (R37CA54281, R01 CA63, P01CA33619, U01CA136792 and U01CA98758). For the Stanford Global Reference Panel, individuals from Puno, Peru were recruited by J. Baker and C. Bustamante, with funding from the Burroughs Welcome Fund, and individuals from Rapa Nui (Easter Island) were recruited by K. Sandoval Mendoza and A. Moreno Estrada, with funding from the Charles Rosenkranz Prize for Health Care Research in Developing Countries. The Women's Health Initiative (WHI) programme is funded by the NHLBI (HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C and HHSN271201100004C). The GALA II study and E. G. Burchard are supported by the Sandler Family Foundation, the American Asthma Foundation, the RWJF Amos Medical Faculty Development Program, the Harry Wm. and Diana V. Hind Distinguished Professor in Pharmaceutical Sciences II, the NHLBI (R01HL117004, R01HL128439, R01HL135156 and X01HL134589), the NIEHS (R01ES015794, R21ES24844), the NIMHD (P60MD006902, R01MD010443, RL5GM118984) and the Tobacco-Related Disease Research Program (24RT-0025). We acknowledge the following GALA II co-investigators for recruitment of individuals, sample processing and quality control: C. Eng, S. Salazar, S. Huntsman, D. Hu, A. C.Y. Mak, L. Caine, S. Thyne, H. J. Farber, P. C. Avila, D. Serebrisky, W. Rodriguez-Cintron, Jose R. Rodriguez-Santana, R. Kumar, L. N. Borrell, E. Brigino-Buenaventura, A. Davis, M. A. LeNoir, K. Meade, S. Sen and F. Lurmann, and we thank the staff and participants who contributed to the GALA II study.

Author contributions I.M.H. conceived and directed the study. D.E.L. and H.J.A. developed the final version of the SV calling pipeline, constructed the SV callsets and performed the data analyses. C.C. and R.M.L. helped design the SV calling pipeline. A.A.R. contributed to long-read validation. I.D. was instrumental in the migration of workflows to the Google Cloud Platform. K.L.K. assisted with data management. E.S.L., B.M.N. and N.O.S. provided input on population genetic analyses. W.J.S., D.M.M., E.S.L., B.M.N., M.C.Z., C.R., T.C.M., S.B., S.K.D., I.M.H. and N.O.S. directed data production, processing and management at their respective sites, and edited the manuscript. Members of the NHGRI CCDG consortium provided samples, produced sequencing data and coordinated and administered data-sharing efforts. H.J.A., D.E.L. and I.M.H. wrote the manuscript.

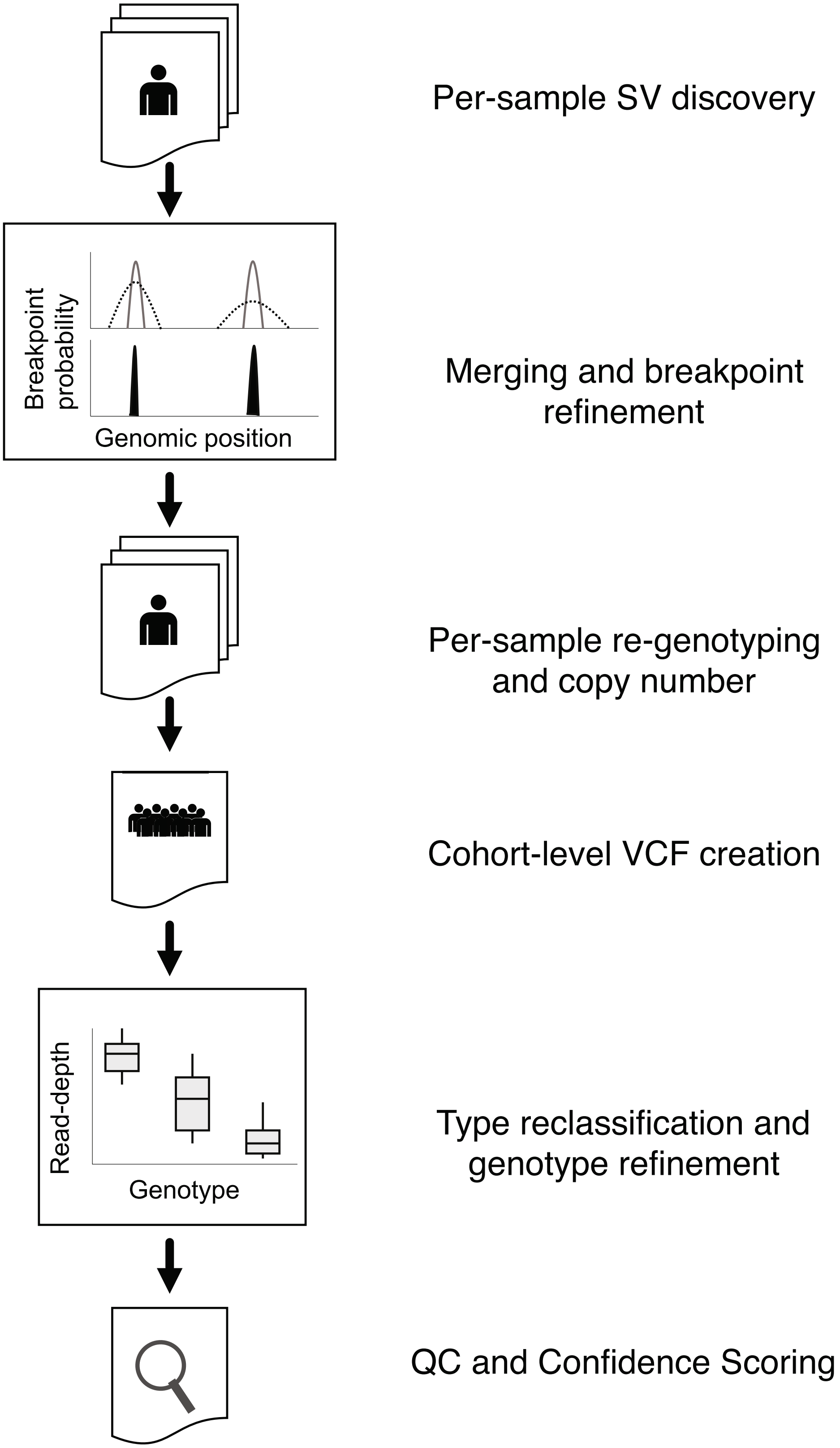
Competing interests The authors declare no competing interests.

Additional information

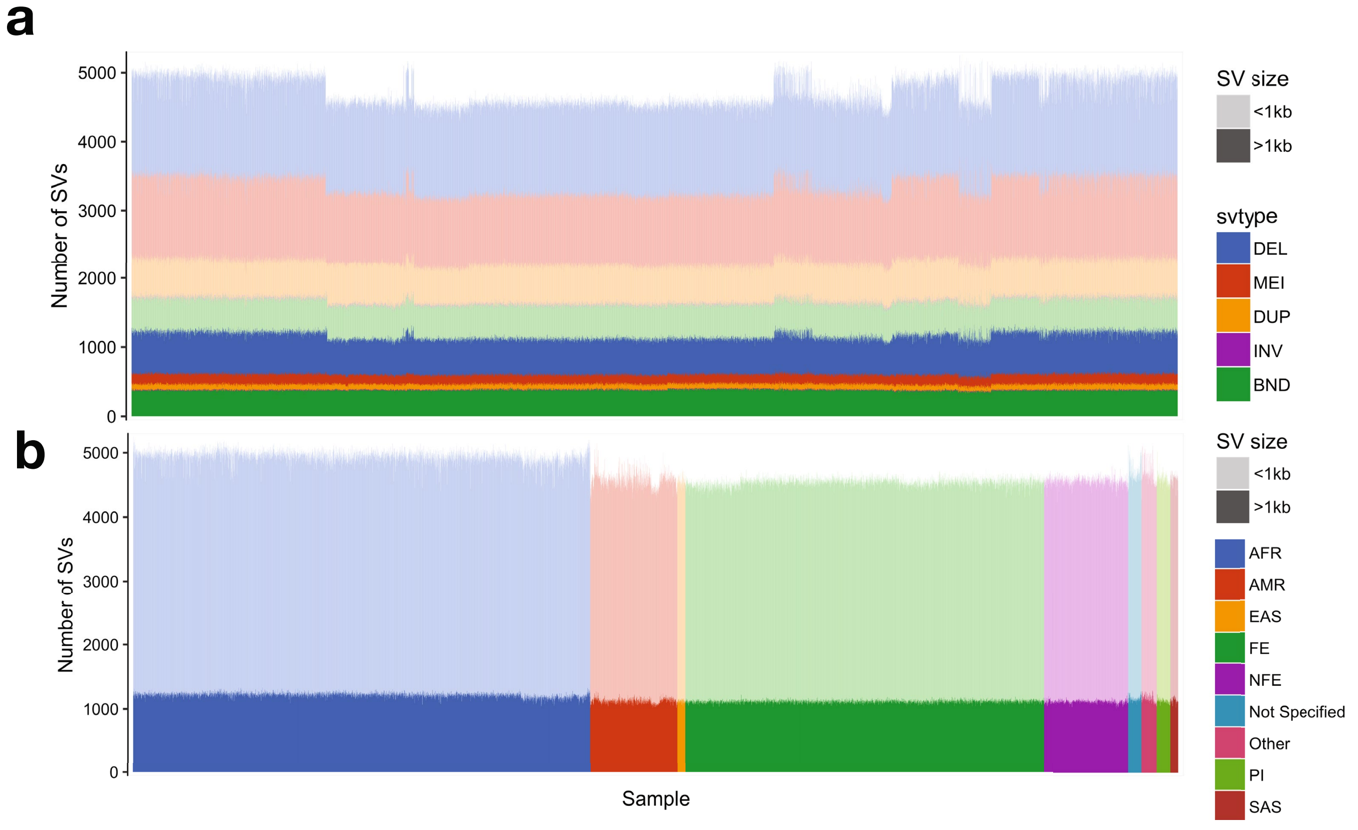
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2371-0.

Correspondence and requests for materials should be addressed to I.M.H.

Reprints and permissions information is available at http://www.nature.com/reprints.

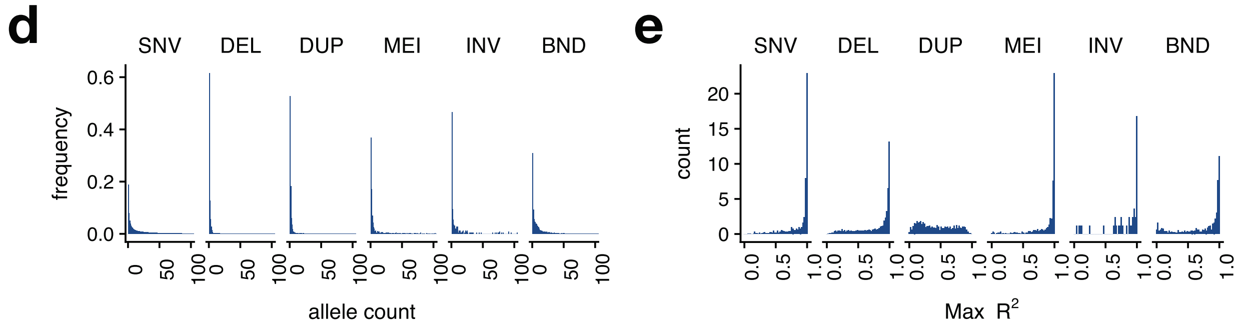


Extended Data Fig. 1 | SV mapping pipeline. SVs are detected within each sample using LUMPY. Breakpoint probability distributions are used to merge and refine the position of detected SVs within a cohort, followed by parallelized re-genotyping and copy-number annotation. Samples are merged into a single cohort-level VCF file, variant types reclassified and genotypes refined with svtools using the combined breakpoint genotype and read-depth information. Finally, sample-level quality control (QC) and variant confidence scoring is conducted to produce the final callset.

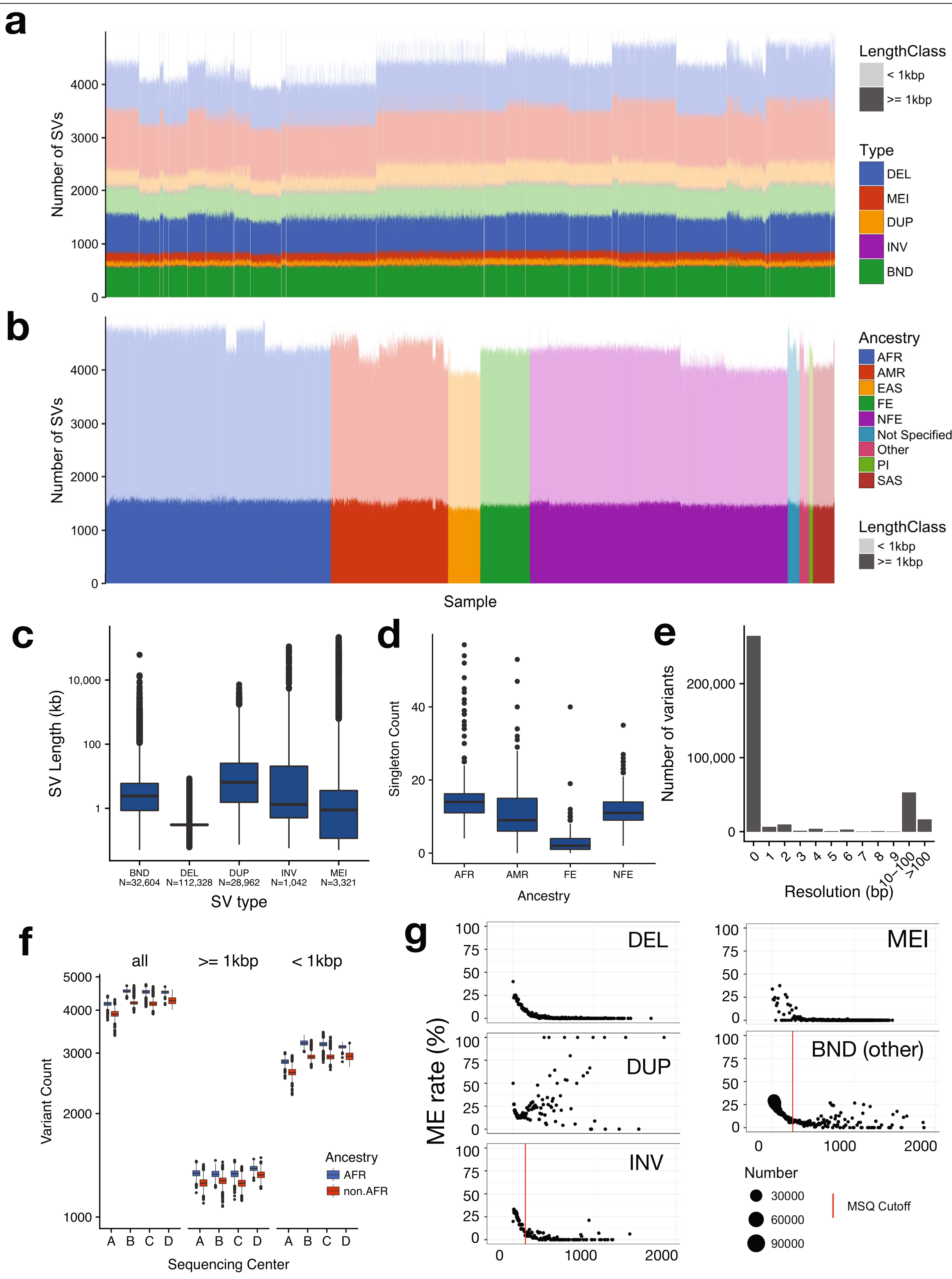


c

Type	Variant counts				Mendelian Error Rate
	Ultra-Rare	Rare	Low Freq	Common	
DEL	49322	25054	2831	4562	0.041
DUP	9056	4276	542	1344	0.113
MEI	621	1923	268	1960	0.009
INV	343	247	35	36	0.009
Low-confidence INV	828	1528	455	874	0.082
BND	7064	7782	465	1242	0.022
Low-confidence BND	29787	87056	20606	21199	0.14



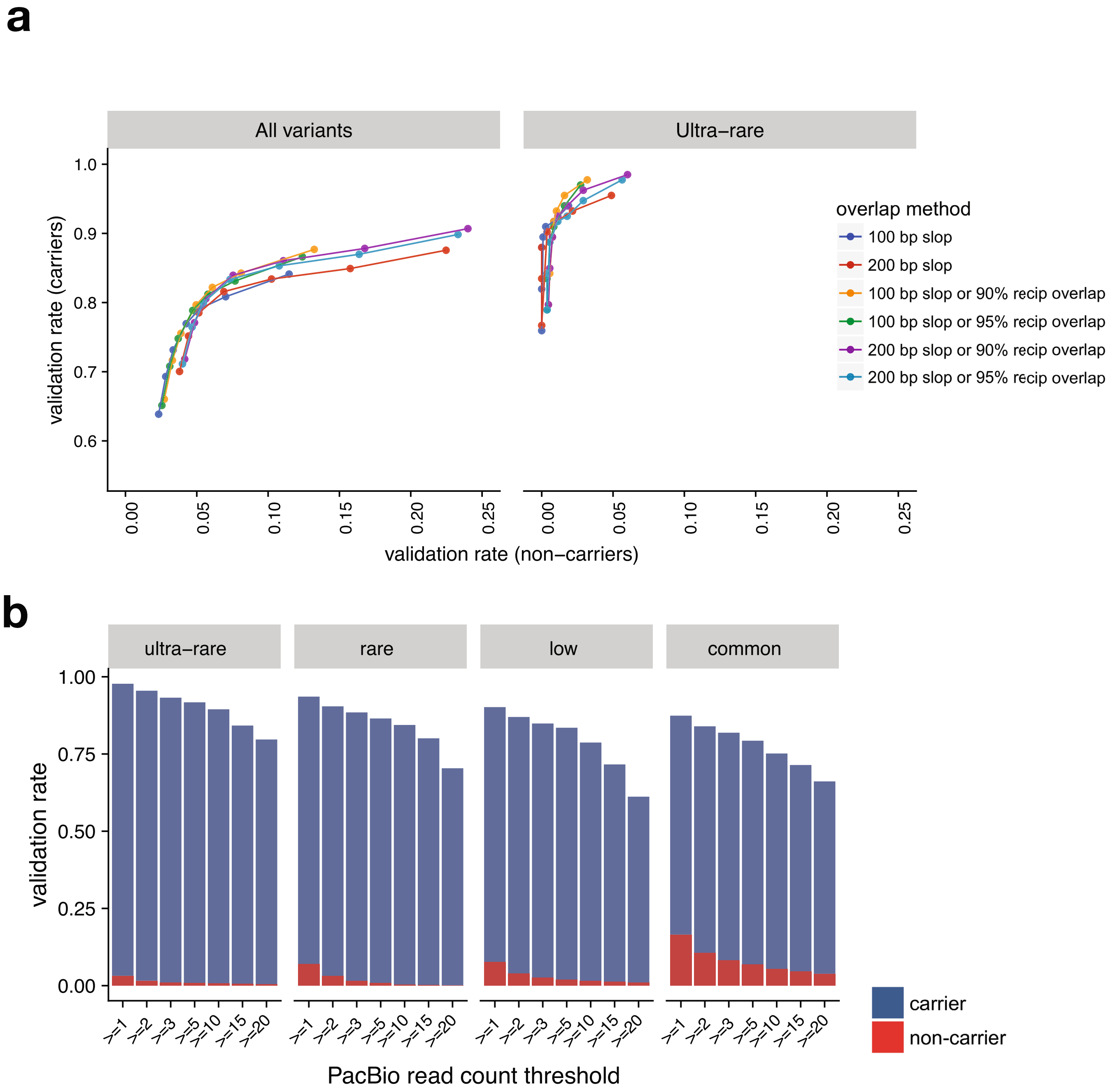
Extended Data Fig. 2 | The B37 callset. a, Variant counts (*y* axis) for each sample (*x* axis) in the callset, ordered by cohort. Large (>1 kb) variants are shown in dark shades and smaller variants in light shades. **b**, Variant counts per sample, ordered by self-reported ancestry according to the colour scheme on the right. Abbreviations as in Fig. 1a. Note that African-ancestry samples show more variant calls, as expected. **c**, Table showing the number of variant calls by variant type and frequency class, and Mendelian error rate by variant type. **d**, Histogram of allele count for each variant class, showing alleles with counts ≤ 100. **e**, Linkage disequilibrium of each variant class as represented by maximum R^2 value to nearby SNVs, for $n = 1{,}581$ samples. Note that these distributions mirror those from our previous SV callset for GTEx[4], which was characterized extensively in the context of expression quantitative trait loci.



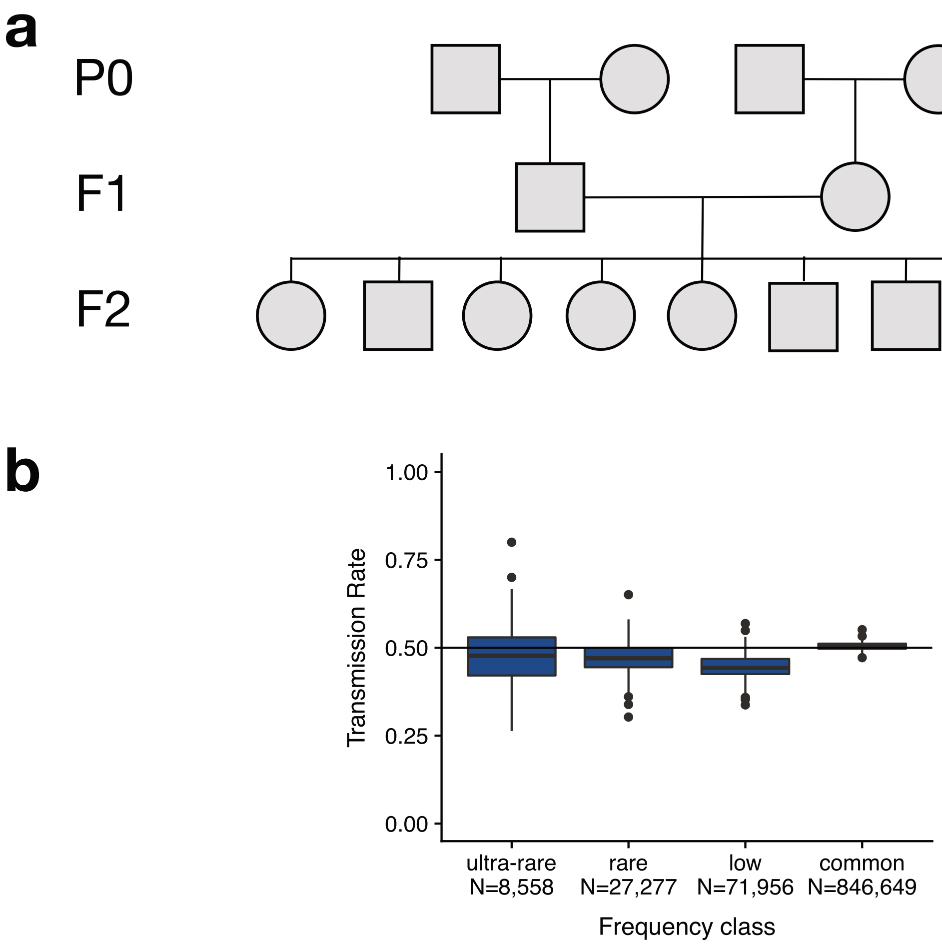
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | The B38 callset. a, Variant counts (*y* axis) for each sample (*x* axis) in the callset, ordered by cohort. Large (>1 kb) variants are shown in dark shades and smaller variants in light shades. **b**, Variant counts per sample, ordered by self-reported ancestry according to the colour scheme on the right. Abbreviations as in Fig. 1a. Note that African-ancestry samples show more variant calls, as expected. Note also that there is some residual variability in variant counts owing to differences in data from each sequencing centre, but that this is mainly limited to small tandem duplications (see **a**), primarily at STRs. **c**, SV length distribution by variant class. **d**, Distribution of the number of singleton SVs detected in samples from different ancestry groups. Only groups with ≥1,000 samples in the B38 callset are shown, and each group was subsampled down to 1,000 individuals before recalculation of the allele frequency. **e**, Histogram showing the resolution of SV breakpoint calls, as defined by the length of the 95% confidence interval of the breakpoint-containing region defined by LUMPY, after cross-sample merging and refinement using svtools. Data are from $n = 360{,}614$ breakpoints, 2 per variant. **f**, Distribution of the number of SVs detected per sample in WGS data from each sequencing centre (*x* axis) for African and non-African (non-AFR) samples, showing all variants (left), and those larger (middle) and smaller (right) than 1 kb in size. Per-centre counts are as follows: centre A, 1,527 AFR, 2,080 non-AFR; centre B, 408 AFR, 2,745 non-AFR; centre C, 2,953 AFR, 2226 non-AFR; centre D, 150 AFR, 2,534 non-AFR. **g**, Plots of Mendelian error (ME) rate (*y* axis) by MSQ for each variant class. Dot size is determined by point density (right) and the threshold used to determine high and low confidence SVs are shown by the vertical lines. All box plots in indicate the median (centre line) and the first and third quartiles (box limits); whiskers extend 1.5 × IQR.

Extended Data Fig. 4 | PCA for the B37 callset. PCA was performed using a linkage disequilibrium-pruned subset of high-confidence DEL and MEI variants, with MAF > 1%. Self-reported ancestry is shown using the colour scheme on the right, with abbreviations as in Fig. 1a.



Extended Data Fig. 5 | Validation of SV calls by PacBio long reads in nine control samples. *n* = 9,905 variants. **a**, Validation rates in variant carriers (*y* axis) versus validation rates in non-carriers (that is, false validations; *x* axis) for each method of determining variant overlap, for a range of supporting-read-count thresholds. Ultra-rare variants (*n* = 133) are shown separately on the right. For each variant overlap method, each data point represents a distinct read-count threshold (≥1, 2, 3, 5, 10, 15 or 20 PacBio reads) that was used to determine validation of SV calls by long-read alignments. Two methods were used for determining overlap between SV coordinates and long-read alignments while accounting for positional uncertainty: (1) BEDTools pairtopair, requiring overlap between the pair of breakpoint intervals predicted by short-read SV mapping and the pair of breakpoint intervals predicted by long-read alignment, allowing 100 bp or 200 bp of 'slop'; and (2) BEDTools intersect, requiring 90% or 95% reciprocal overlap between the coordinates spanned by the SV predicted by short-read SV mapping and the SV predicted by long-read alignment. Here, we plot the first criteria by themselves, and in pairwise combination with the latter (see key on the right of the figure). Note that Supplementary Table 3 is based on the '100 bp slop or 90% reciprocal overlap' method, requiring at least two PacBio reads. **b**, Validation rates by frequency class for variant carriers and non-carriers with increasing PacBio supporting-read thresholds, shown using the same overlap method as in Supplementary Table 3. Variant counts per frequency class are as follows: ultra-rare, *n* = 133; rare, *n* = 734; low frequency, *n* = 1,361; common, *n* = 7,677.



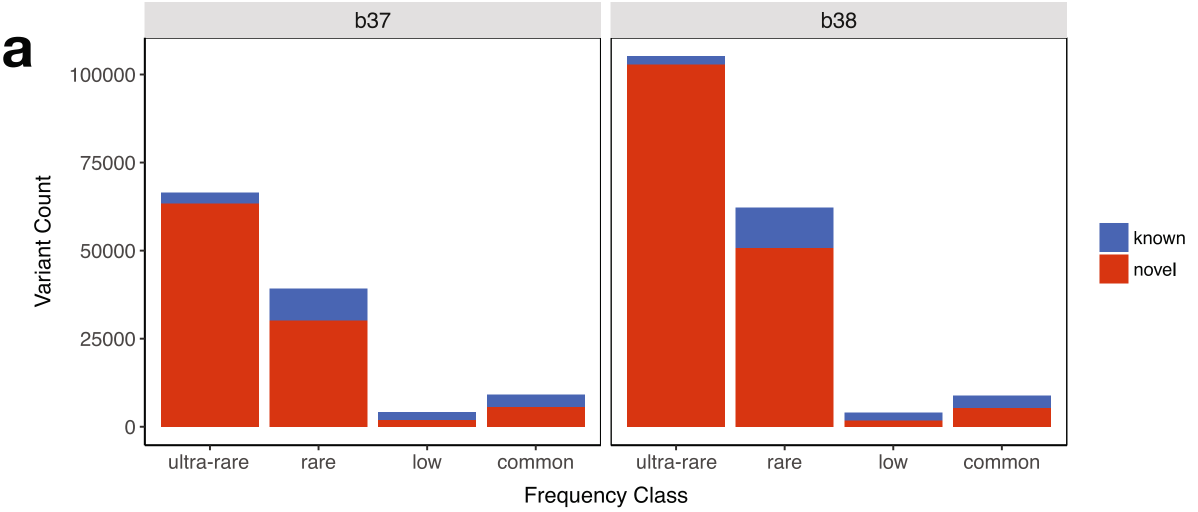
c

Frequency bin	total count	Mendelian errors	Mendelian error rate
common	1,446,555	45,553	0.031
low	31,612	2,927	0.093
rare	11,105	480	0.043
ultra-rare	3,441	82	0.024

d

SV type	Mendelian errors	Total count	Mendelian error rate
DEL	55	2,291	0.024
DUP	21	495	0.042
INV	0	7	0.000
MEI	4	4	1.000
BND	2	644	0.003
All	82	3,441	0.024

Extended Data Fig. 6 | Mendelian inheritance analysis in a set of three-generation CEPH pedigrees comprising 409 parent–offspring trios. **a**, Example structure of a single CEPH pedigree indicating nomenclature of the parental (P_0), first (F_1) and second (F_2) generations. **b**, Transmission rate of SVs from different allele-frequency classes including SVs that are unique to a single family (ultra-rare), rare (<1%), low frequency ('low'; 1–5%) and common (>5%). **c**, Table showing the number and rate of Mendelian errors by allele-frequency class. **d**, Table showing the number and rate of Mendelian errors for SVs that are unique to a single family, for each SV type.

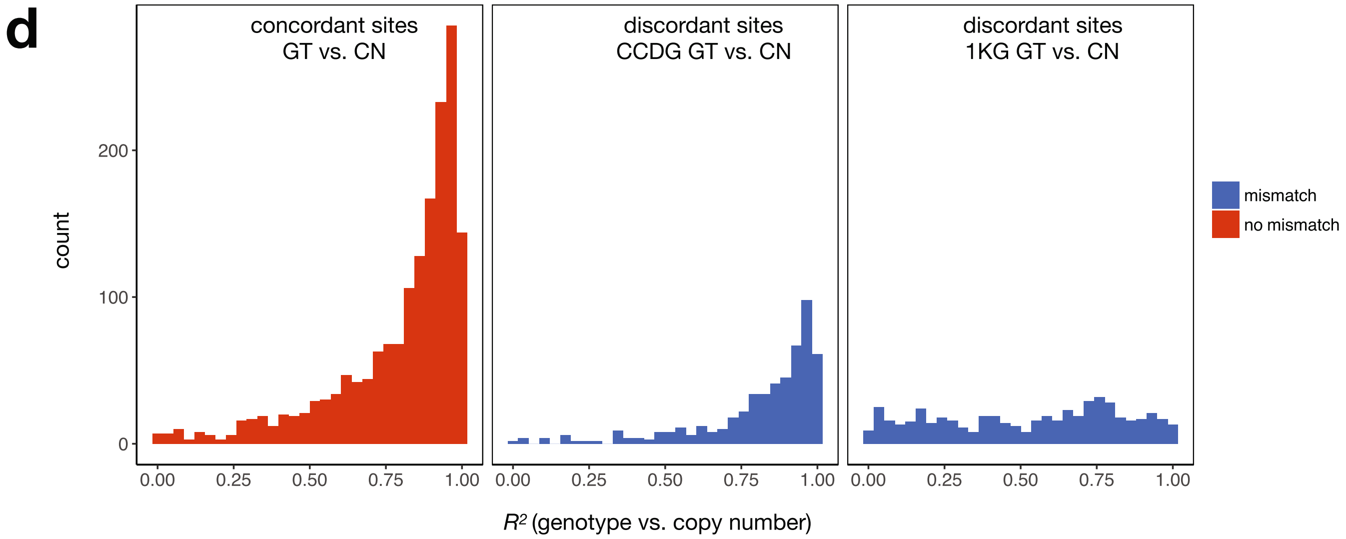


b

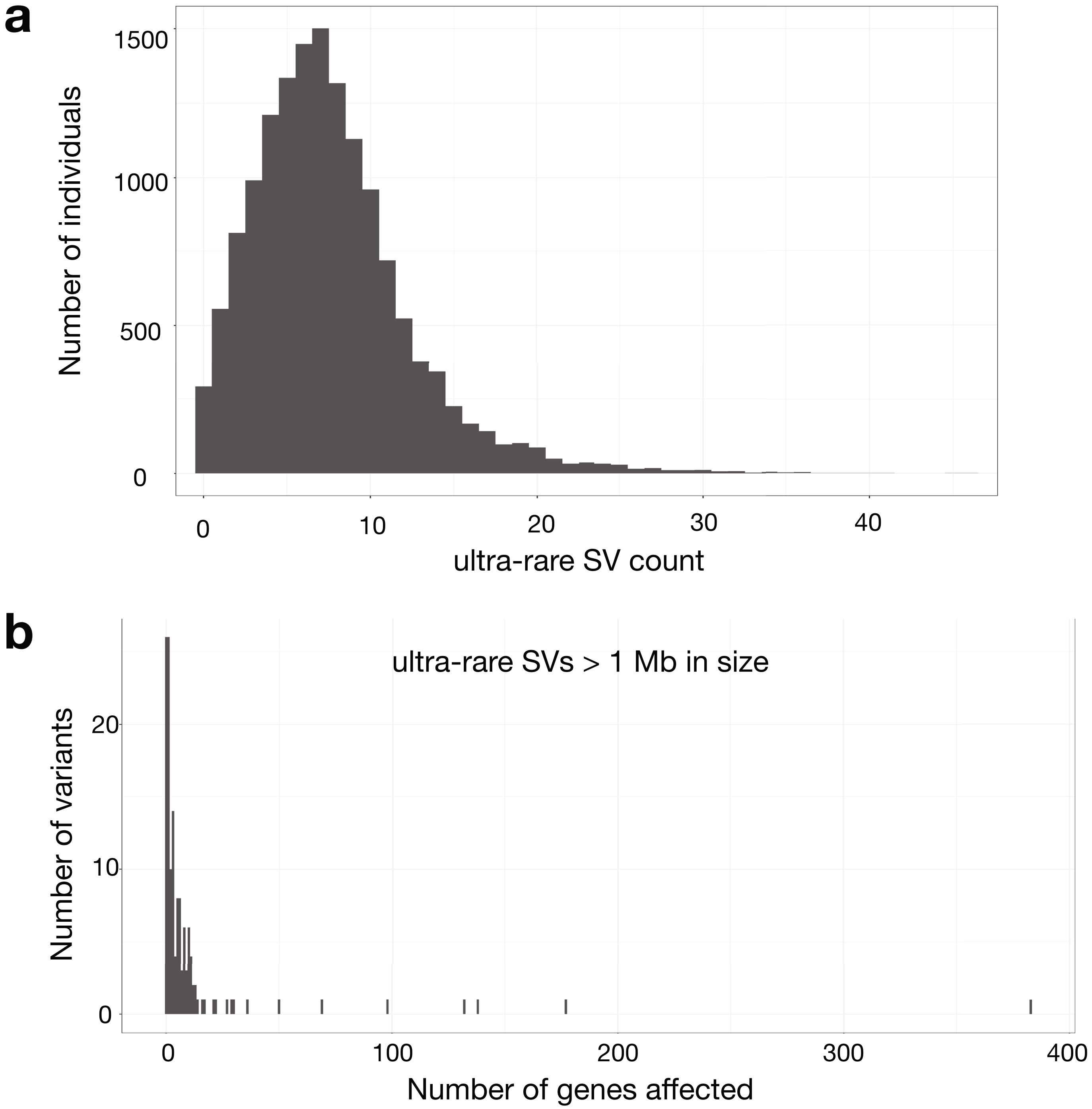
		1000 Genomes GT		
		0/0	0/1	1/1
B38 GT	0/0	6283	42	1
	0/1	138	3813	29
	1/1	15	992	1888

c

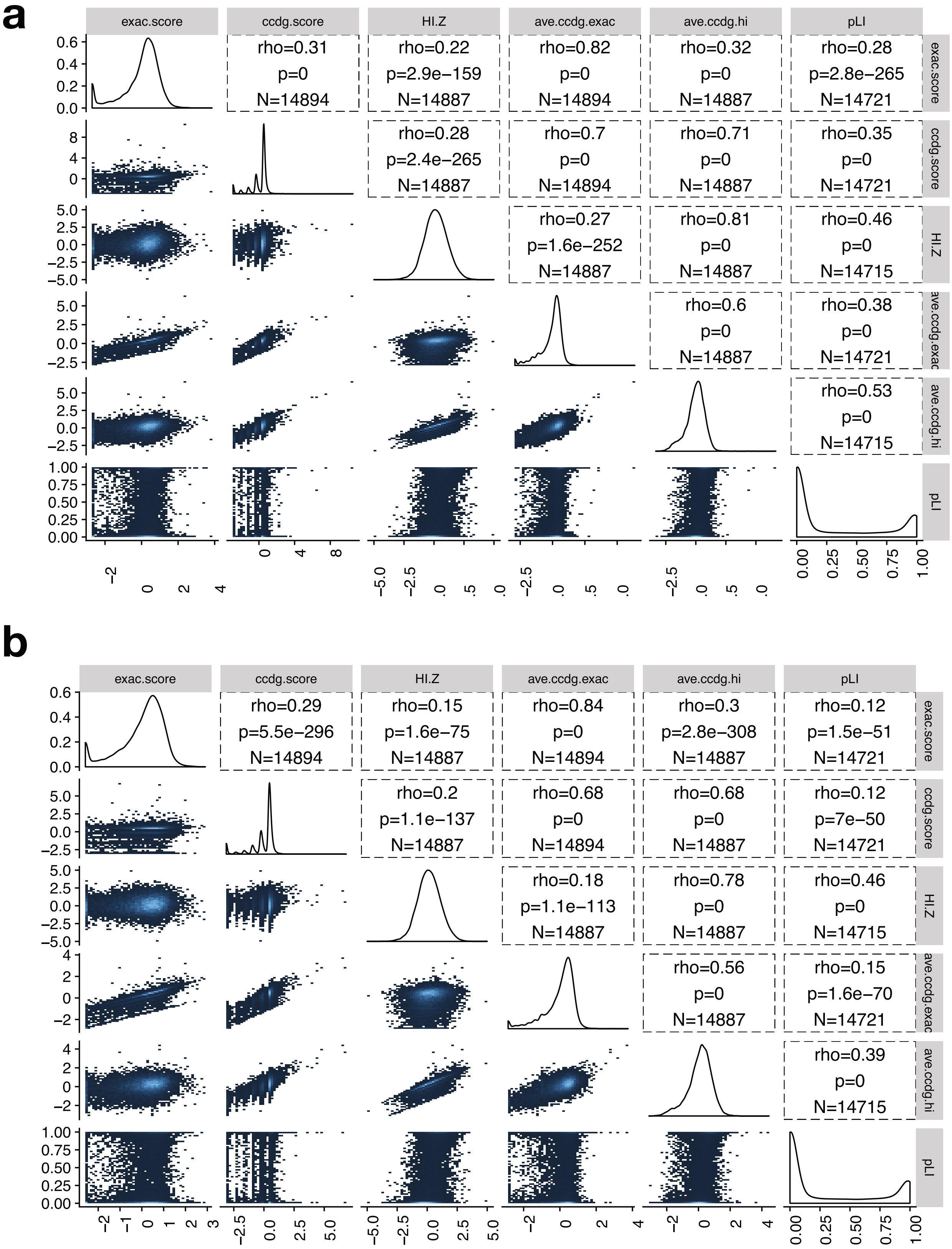
svtype	total	kappa	frac concordant
DEL	7150	0.964	0.981
BND	1495	0.879	0.929
MEI	4556	0.685	0.787
all	13201	0.852	0.908



Extended Data Fig. 7 | Comparison of SV calls and genotypes to the 1KG phase 3 callset. a, Number of known and novel SVs in the B37 (left) and B38 (right) callsets, shown by frequency class. **b**, Table showing the genotypes (GT) reported in our B38 callset[5] (rows) versus the 1KG callset (columns) at SVs identified by both studies among the five samples included in both callsets. **c**, Table showing genotype concordance by SV type including the fraction of concordant calls and Cohen's κ coefficient. **d**, Distribution of correlation (R^2) between genotype information determined by breakpoint-spanning reads and estimates of copy number (CN) determined by read-depth analysis for the SVs shown in **b**, **c** when genotype information between the B38 and the 1KG callset is concordant (left) or discordant (middle, right). At sites with discordant genotypes, correlation with copy-number information is typically higher for genotypes from the B38 callset (middle) than the 1KG callset (right).



Extended Data Fig. 8 | Ultra-rare SVs in the B38 callset. *n* = 14,623. **a**, Histogram showing the number of ultra-rare SVs per individual (ultra-rare is defined as singleton variants private to a single individual or nuclear family). **b**, Histogram showing the number of genes affected by ultra-rare SVs larger than 1 Mb in size.



Extended Data Fig. 9 | Correlations between dosage sensitivity scores for CNV in the combined callset. *n* = 17,795. **a**, Results for deletion variants. The ExAC score is the published ExAC DEL intolerance score[45]; the CCDG score is similarly calculated from our data, using CCDG deletions; pLI is the published loss-of-function intolerance score from ExAC[27]; 'HI.Z' is the negative of the inverse-normal transformed haploinsufficiency score from DECIPHER[46]; 'ave.ccdg.exac' is the arithmetic mean of the CCDG and ExAC DEL intolerance scores; and 'ave.ccdg.hi' is the arithmetic mean of the CCDG and HI-Z scores. The correlations shown are Spearman rank correlations (rho); *P* values are calculated by two-sided Spearman rank correlation test; and *N* represents the number of genes included in the test. **b**, Results for duplication variants, using the same naming conventions as in **a**.

Extended Data Table 1 | Ancestry, ethnicity and continental origin of the samples analysed in this study

a

Ancestry	Build 37	Build 38	Combined
AFR	3683	5501	6170
AMR	698	4165	4186
EAS	65	929	972
FE	2898	1207	2884
NFE	682	9588	10254
Not Specified	105	428	436
Other	123	751	777
PI	110	87	110
SAS	62	519	558

b

Ethnicity	Build 37	Build 38	Combined
Hispanic	586	2829	2829
Non-Hispanic	3758	8022	10559
Not Specified	4082	12324	12959

c

Continent	Build 37	Build 38	Combined
African	66	24	66
Asian	32	1272	1272
Caribbean	279	1815	1815
East Asian	43	0	43
European	2985	1219	2971
North American	4641	18563	19800
Oceanic	41	18	41
Central Asian/Siberian	26	0	26
South American	274	264	274
South Asian	39	0	39

For each table, the number of samples in the B37 and B38 callsets are shown separately, and the non-redundant combined set is shown on the right. Abbreviations as in Fig. 1a.

nature research

Corresponding author(s): Hall

Last updated by author(s): Apr 4, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	no software was used
Data analysis	All sequence data were aligned and processed as described in the methods section. For the 'b37' callset, data were processed using the speedseq pipline. For the 'b38' callset data were processed according to the functional equivalence standard. We used LUMPY (v0.2.13) for per-sample SV calling followed by cohort-level merging, re-genotyping, etc, using the svtools (v0.3.2) workflow as detailed in the Methods section to produce a joint, cohort-level vcf. Dataset qc was performed using bcftools (v1.3.1) and vawk (https://github.com/cc2qe/vawk). The SNV/indel callset was produced using GATK HaplotypeCaller (v3.5-0-g36282e4) as detailed in the methods and annotated using vep and LOFTEE (v0.2.2-beta). Validation of SV by PacBio long reads was performed using custom code in (https://github.com/abelhj/long-read-validation/tree/master). All further analysese were performed using bedtools (v2.23.0) and R (v3.3.3).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The sequencing data can be accessed through dbGaP (https://www.ncbi.nlm.nih.gov/gap) under accession numbers provided in Supplemental Table 7. PacBio long read data used for SV validation can be accessed through SRA, under accession numbers provided in Supplemental Table 2. The set of high-confidence HGSVC long-read derived SV calls, validated by our independent PacBio data and used as a truth set can be found in Supplementary_File4.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size was determined based on the number of distinct individuals in the callsets.
Data exclusions	As detailed in the Methods sections, samples with per-sample variant counts of any type exceeding the median+6*MAD were excluded (per our standard qc practice). A set of 64 samples were excluded because they appeared to be duplicates (or monozygotic twins) of other samples in the callset. (One per duplicate pair was excluded at random.) Additional samples were excluded because we could not obtain consent for aggregate sharing. (See methods for details.)
Replication	This was an observational study, there was no attempt at replication.
Randomization	This was an observational study, there was no randomization.
Blinding	This was an observational study, there was no blinding.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Whole-genome sequencing of patients with rare diseases in a national health system

https://doi.org/10.1038/s41586-020-2434-2

Received: 23 December 2018

Accepted: 5 May 2020

Ernest Turro[1,2,3]✉, William J. Astle[3,4], Karyn Megy[1,2], Stefan Gräf[1,2,5], Daniel Greene[1,3],
Olga Shamardina[1,2], Hana Lango Allen[1,2], Alba Sanchis-Juan[1,2], Mattia Frontini[1,4,6], Chantal Thys[7],
Jonathan Stephens[1,2], Rutendo Mapeta[1,2], Oliver S. Burren[5,8], Kate Downes[1,2],
Matthias Haimel[1,2,5], Salih Tuna[1,2], Sri V. V. Deevi[1,2], Timothy J. Aitman[9,10], David L. Bennett[11,12],
Paul Calleja[13], Keren Carss[1,2], Mark J. Caulfield[14,15], Patrick F. Chinnery[2,16,17], Peter H. Dixon[18],
Daniel P. Gale[19,20], Roger James[1,2], Ania Koziell[21,22], Michael A. Laffan[23,24], Adam P. Levine[19],
Eamonn R. Maher[25,26,27], Hugh S. Markus[28], Joannella Morales[29], Nicholas W. Morrell[2,5],
Andrew D. Mumford[30,31], Elizabeth Ormondroyd[12,32], Stuart Rankin[13], Augusto Rendon[1,14],
Sylvia Richardson[3], Irene Roberts[12,33,34], Noemi B. A. Roy[12,33,35], Moin A. Saleem[36,37],
Kenneth G. C. Smith[5,8], Hannah Stark[2,38], Rhea Y. Y. Tan[28], Andreas C. Themistocleous[11],
Adrian J. Thrasher[39], Hugh Watkins[32,35,40], Andrew R. Webster[41,42], Martin R. Wilkins[43],
Catherine Williamson[18,44], James Whitworth[25,26,27], Sean Humphray[45], David R. Bentley[45],
NIHR BioResource for the 100,000 Genomes Project*, Nathalie Kingston[1,2], Neil Walker[1,2],
John R. Bradley[2,5,26,46,47], Sofie Ashford[2,38], Christopher J. Penkett[1,2], Kathleen Freson[7],
Kathleen E. Stirrups[1,2], F. Lucy Raymond[2,25]✉ & Willem H. Ouwehand[1,2,4,6,48]✉

Most patients with rare diseases do not receive a molecular diagnosis and the aetiological variants and causative genes for more than half such disorders remain to be discovered[1]. Here we used whole-genome sequencing (WGS) in a national health system to streamline diagnosis and to discover unknown aetiological variants in the coding and non-coding regions of the genome. We generated WGS data for 13,037 participants, of whom 9,802 had a rare disease, and provided a genetic diagnosis to 1,138 of the 7,065 extensively phenotyped participants. We identified 95 Mendelian associations between genes and rare diseases, of which 11 have been discovered since 2015 and at least 79 are confirmed to be aetiological. By generating WGS data of UK Biobank participants[2], we found that rare alleles can explain the presence of some individuals in the tails of a quantitative trait for red blood cells. Finally, we identified four novel non-coding variants that cause disease through the disruption of transcription of *ARPC1B*, *GATA1*, *LRBA* and *MPL*. Our study demonstrates a synergy by using WGS for diagnosis and aetiological discovery in routine healthcare.

Rare diseases affect approximately 1 in 20 people, but only a minority of patients receive a genetic diagnosis[3]. Approximately 10,000 rare diseases are known, but fewer than half have a resolved genetic aetiology[1]. Even for diseases with a resolved aetiology, the prospects for diagnosis are severely diminished by fragmentary phenotyping and the restriction of testing to disease-specific panels of genes. It may require more than 20 physician visits over several years to determine a molecular cause[4]. Recent development of WGS technology enables systematic, comprehensive genetic testing in integrated health systems, together with aetiological discovery in the coding and non-coding genome.

We performed WGS for 13,037 individuals enrolled at 57 National Health Service (NHS) hospitals in the United Kingdom and 26 hospitals in other countries (Fig. 1a, Extended Data Fig. 1a and Supplementary Table 1), in three batches, to clinical standard (Fig. 1b). The participants were distributed approximately uniformly across the sexes (Supplementary Table 1) and approximately according to the distribution reported by the UK census across ethnic groups (Fig. 1c; https://www.ons.gov.uk/census/2011census). Each participant was assigned to one of 18 domains with pre-specified enrolment criteria (Supplementary Table 1): 7,388 individuals were assigned to one of 15 rare disease domains, 50 individuals to a control domain, 4,835 individuals to a domain called the Rare Diseases Pilot of Genomics England Ltd (GEL) and 764 individuals to a domain comprising UK Biobank participants with extreme red blood cell indices (Extended Data Fig. 1b, Supplementary Information and Supplementary Table 1). Sample sizes varied across domains, primarily owing to differences in recruitment rates, limiting the efficiency of the study design. In total, 9,802 of the participants (75%) had a rare disease or an extreme measurement of a quantitative trait, of whom 9,024 were probands and 778 were affected relatives. The patients presented with pathologies of many organ systems, which we phenotyped using Human Phenotype Ontology (HPO) terms for all of the rare disease domains except the domain comprising Leber's hereditary optic neuropathy and the domain comprising Ehler–Danlos/Ehler–Danlos-like syndromes (Fig. 2a and Extended Data

---

A list of affiliations appears at the end of the paper.

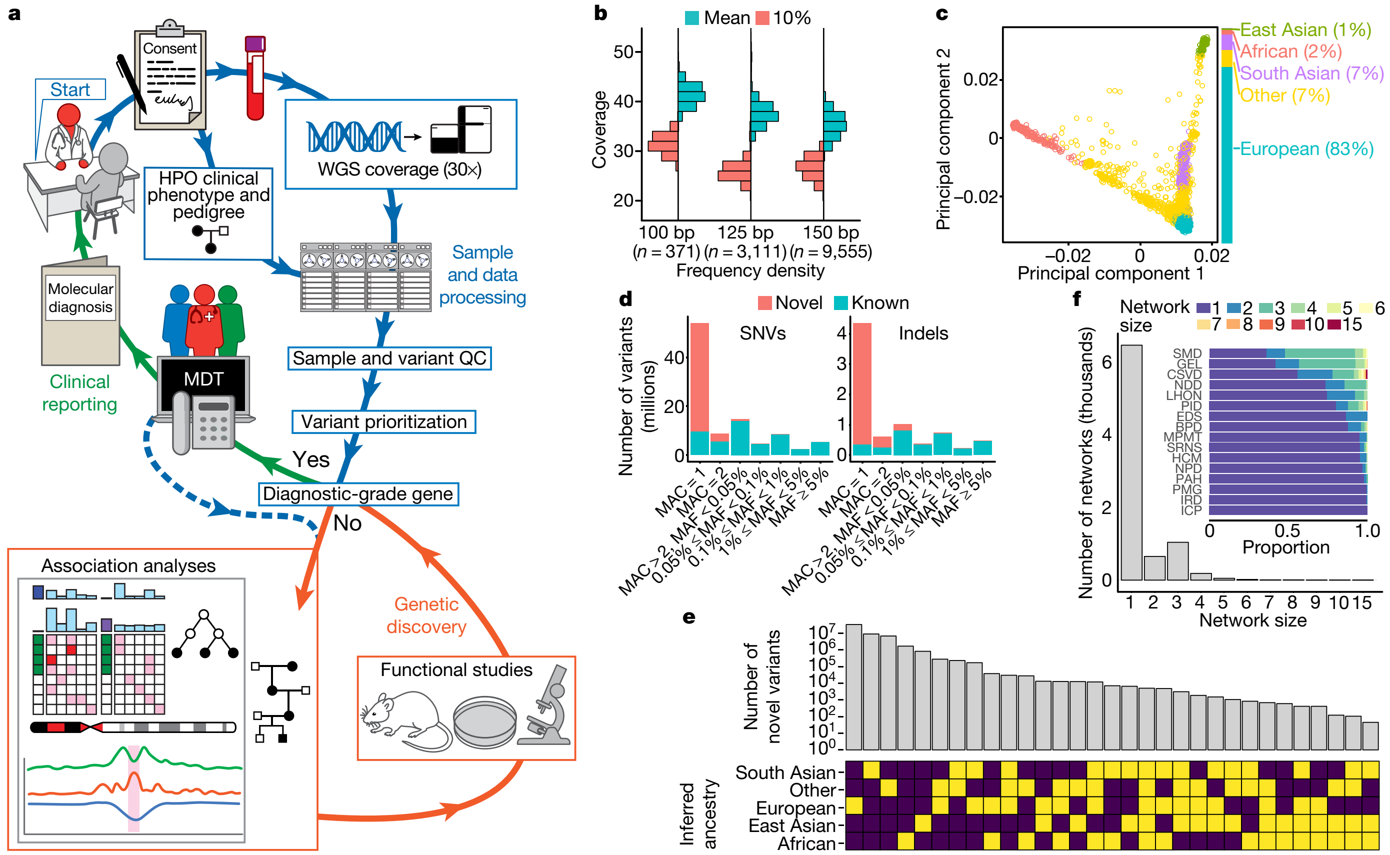


Fig. 1 | Study overview. a, Schematic of the diagnostic and research processes. Blue, patients are recruited, HPO and pedigree data are collected, DNA is extracted and sequenced and WGS data are transferred for quality control and variant prioritization. Green, variants are assessed and diagnoses are returned. Orange, the complete data are analysed by association and co-segregation to identify aetiological variants, disease-mediating genes and regulatory regions; functional studies and model systems are used to study disease mechanisms. **b**, Histograms of read coverage across the 13,037 participants, stratified by WGS read length (100 bp, 125 bp and 150 bp). **c**, Projection of genetic data of the 13,037 participants onto the first two principal components of variation in the 1000 Genomes Project and the distribution of participant ancestry. **d**, Histograms illustrating the observed distribution of the minor allele frequency (MAF) of variants called in the MSUP (*n* = 10,259), stratified by type (SNV or indel). Variants are labelled novel if they were uncatalogued in the 1000 Genomes, UK10K, TOPMed, gnomAD and HGMD Pro databases. MAC, minor allele count. **e**, The number of novel variants stratified by the ancestry groups in which they were observed (yellow, present; navy, absent). **f**, The sizes of genetically determined networks of closely related individuals across the 13,037 participants. Inset, distributions of network sizes for each rare disease domain. BPD, bleeding, thrombotic and platelet disorders; CSVD, cerebral small vessel disease; EDS, Ehler–Danlos and Ehler–Danlos-like syndromes; HCM, hypertrophic cardiomyopathy; ICP, intrahepatic cholestasis of pregnancy; IRD, inherited retinal disorders; LHON, Leber's hereditary optic neuropathy; MPMT, multiple primary malignant tumours; NDD, neurological and developmental disorders; NPD, neuropathic pain disorders; PAH, pulmonary arterial hypertension; PID, primary immune disorders; PMG: primary membranoproliferative glomerulonephritis; SMD, stem cell and myeloid disorders; SRNS, steroid-resistant nephrotic syndrome.

Fig. 1c). The GEL domain released only a binary affection phenotype for these analyses. In total, 19,605 HPO terms were assigned to patients.

Following bioinformatic analysis (Extended Data Fig. 2–4), we considered a maximal set of 10,259 unrelated participants (MSUP), in which we identified 172,005,610 short variants. These variants comprised 157,411,228 (91.5%) single-nucleotide variants (SNVs) and 14,594,382 (8.5%) small insertions or deletions (indels) of ≤50 base pairs (bp) (Extended Data Fig. 5). Of these SNVs and indels, 48.6% and 40.8%, respectively, were absent from major public variant databases (Fig. 1d) and 54.8% had a minor-allele count of 1. Of these singleton variants, 82.6% were novel. Only 9.08% of the novel variants had a minor-allele count > 1; in these cases, the minor allele was typically carried exclusively by individuals with similar population ancestry (Fig. 1e). SNVs and indels were well represented in major variant databases if they were common in our dataset; however, consistent with theory, most variants were very rare and, of these, most were uncatalogued. We called 177,550 distinct large deletions (>50 bp) across the 13,037 participants by synthesizing inferences from two algorithms. We also called more complicated types of structural variant, such as inversions; however, this was unreliable and we could not reconcile the calls across individuals (Supplementary Information). Only 13 (0.1%) individuals had non-standard WGS-determined sex chromosomal karyotypes (Extended Data Fig. 3e–g). We inferred familial relationships from the genetic data (Supplementary Information). Owing to the enrolment strategies, most families were singletons (Fig. 1f).

Clinical reporting

For each of the 15 rare disease domains, we reviewed the scientific literature to establish a list of diagnostic-grade genes (DGGs) and to identify the corresponding transcripts (Supplementary Information). The lists ranged in length from two for the intrahepatic cholestasis of pregnancy domain to 1,423 for the neurological and developmental disorders domain. The lists were not mutually exclusive because mutations in some genes cause pathologies that were compatible with the enrolment criteria of multiple domains (Fig. 2b). Twelve multidisciplinary teams (MDTs) with domain-specific expertise examined the rare variants observed in DGGs in the context of the HPO phenotypes. They

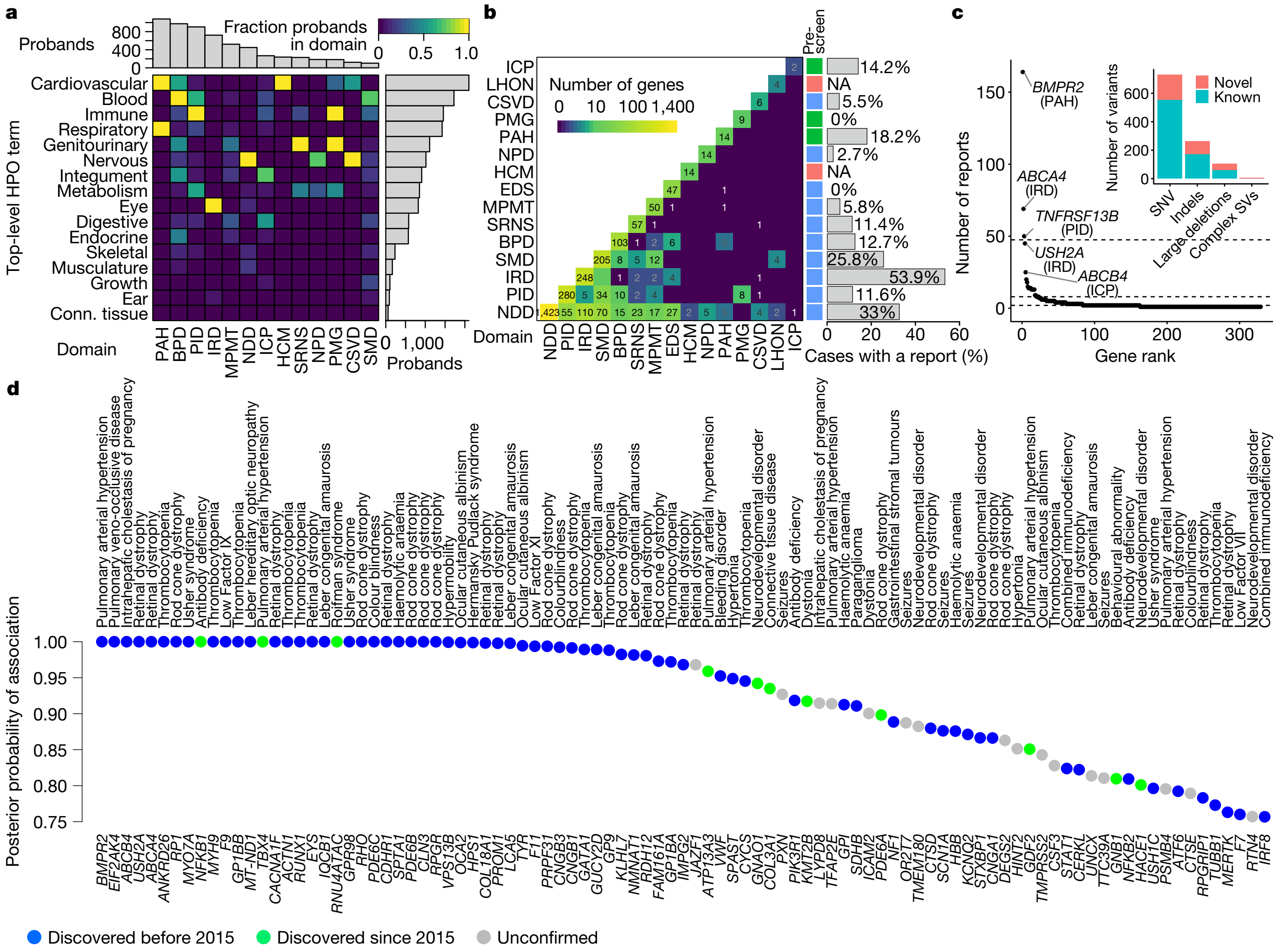


Fig. 2 | Variant reporting and genetic associations with rare diseases. a, The frequency of probands by domain (top) and by top-level HPO-phenotype abnormality term (right). The heat map shows the proportion of probands in each domain that were assigned a particular top-level HPO term (shown abbreviated). **b**, The number of DGGs shared by pairs of domains (left). Pre-screening level for each domain indicated in red (full), blue (partial) or green (none). The proportion of cases for which a clinical report was issued (right). **c**, The number of reports issued by DGG ordered inversely by count. Dashed lines indicate quartiles of the count distribution. Inset, the number of distinct clinically reported variants stratified by variant type. The colours in each bar indicate the proportion of variants that are known or novel (as defined in the main text). **d**, BeviMed posterior probabilities for genetic association >0.75. The colours indicate whether the associations were established in the scientific literature before 2015, since 2015 or remain unconfirmed. *GPR98* is also known as *ADGRV1*; *TMEM180* is also known as *MFSD13A*.

categorized a subset of the variants as 'pathogenic' or 'likely pathogenic' following standard guidelines[5] and assessed their allelic contribution to disease as 'full' or 'partial'. The contribution of a variant was assessed to be full if it was considered to be the only variant for which an isolated reduction in copy number from conception would have eliminated the disease phenotype, otherwise the contribution was assessed to be partial. Clinical reports—which contained molecular diagnoses comprising 1,103 distinct causal variants (731 SNVs, 264 indels, 102 large deletions and 6 complex structural variants) that affected 329 DGGs (Supplementary Table 2)—were issued for 1,138 of the 7,065 (16.1%) patients reviewed. We classed 266 of the 995 SNVs and indels (26.7%) as novel, because they were absent from the Human Gene Mutation Database (HGMD) and were not among the variants in ClinVar with at least one pathogenic or likely pathogenic interpretation and no benign interpretation. We ranked the 329 DGGs by the number of clinical reports in which they featured. The top three DGGs (*BMPR2*, *ABCA4* and *TNFRSF13B*) featured in a quarter of all reports. The subsequent 19 DGGs featured in a further quarter of reports. The remaining 307 DGGs mostly featured in a single report (Fig. 2c and Extended Data Fig. 6). The diagnostic yield by domain ranged from 0% (0 out of 184) of patients for the primary membranoproliferative glomerulonephritis domain to 53.9% (391 out of 725) of patients for the inherited retinal disease domain (Fig. 2b). The variability in diagnostic yield is attributable to heterogeneity in: phenotypic and genetic pre-screening before enrolment, the genetic architecture of the diseases and prior knowledge of genetic aetiologies.

Clinical reporting was enhanced by the use of PCR-free WGS with a mean autosomal depth greater than 35× instead of whole-exome sequencing (WES). For example, we identified a causal SNV encoding a start loss of *HPS6* in a case with Hermansky–Pudlak syndrome that was previously missed by WES. We compared the read coverage of WGS to that of research WES of participants in the UK Biobank[6], INTERVAL[7] and the Columbia University exome-sequencing study for chronic kidney disease (Supplementary Information). Although less costly to generate per sample, the variation in coverage within and between genomic sites that contain known pathogenic SNVs or indels was much greater for WES than WGS (Extended Data Fig. 7). Of the 938 distinct autosomal SNVs reported in this study, the number of autosomal SNVs with

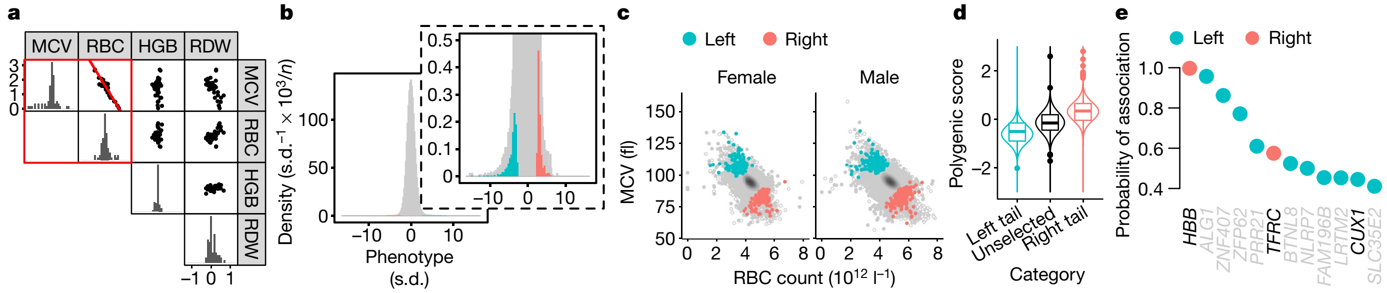


Fig. 3 | Genetic associations with the tails of an RBC trait. a, The distribution of the additive effects of 65 RBC GWAS variants (MAF < 1%) on four RBC traits (acronyms are defined in the Supplementary Information). The red square indicates the bivariate distribution used to develop the selection phenotype. The red line was estimated by Deming regression. **b**, The (standardized) distribution of the selection phenotype (panels showing different *y*-axis ranges) in post-menopausal female and male participants of European ancestry in the UK Biobank without record of illness or treatment that is known to perturb RBC indices (grey) and selected for WGS (turquoise and salmon). The scale of the *x*-axis shows the standard deviations (s.d.) of the phenotype and the scale of the *y*-axis is such that when the units of the axes are disregarded the area of the histogram represents the number of contributing individuals in thousands, where $n = 316{,}739$. Many participants in the tails were unselected (Supplementary Information). **c**, The distribution of RBC count and mean cell volume (MCV) in post-menopausal female (left) and male (right) participants in the UK Biobank. The ellipsoids are contours of kernel density estimates. Open circles, participants ineligible for selection. Non-European ancestry thalassaemias may explain the concentration with high RBC count/low mean cell volume. Coloured circles, participants who have WGS data. **d**, The distribution of a polygenic score for the selection phenotype in the 382 and 368 individuals of European ancestry selected from the left and right tails, respectively, and in 522 European participants in domains other than the UK Biobank (extreme red blood cell traits) domain with pathology explained by rare variants (unselected). The centre mark, lower and upper hinges of the boxplots, respectively, indicate the median, 25th and 75th percentiles. Outliers beyond 1.5× the interquartile range from each hinge are shown. The violin plots show the expected distribution of the polygenic score under a Gaussian variance components model, conditional on the proportion of phenotypic variance explained by the score and the tail-selection thresholds. **e**, BeviMed posterior probabilities for genetic association of each tail (distinguished by colour), for genes with posterior probabilities >0.4. Indicated genes (black font) have strong concordant biological evidence.

insufficient coverage in WES analyses for reliable genotyping ranged between 25 and 99 (2.67–10.5%) across WES datasets (Extended Data Fig. 7). Moreover, deletions that span only a few short exons or part of a single exon are not reliably called by WES[8,9]. Of the 102 distinct large deletions that we reported (length range, 203 bp–16.80 Mb; mean, 786.33 kb; median, 15.91 kb), 22 (21.6%) overlapped only one exon. Although clinical and research WES may have different coverage characteristics, we were unable to obtain an example clinical dataset for comparison.

Measurement of quantitative intermediate phenotypes can elucidate the genetic aetiology in difficult-to-diagnose patients. We considered patients with a clinically determined absence of a protein encoded by a DGG for whom we had called only one explanatory allele and examined the corresponding WGS read alignments for evidence of a variant in compound heterozygosity. Two patients with a severe unexplained bleeding disorder owing to the absence of αIIbβ3 integrin on their platelet membranes carried complex variants in intron 9 of *ITGB3*: one carried a tandem repeat and the other a SINE-VNTR-Alu (SVA) retrotransposon that was not called by structural variant callers, but which generated an excess of improperly mapped reads and was confirmed by long-read sequencing (Extended Data Fig. 8a–e). A third patient had severe haemolytic anaemia owing to absence of the RhD and RhCE proteins in the membranes of her red blood cells, which was caused by a large tandem repeat in *RHAG* (Extended Data Fig. 8f).

Research findings from this study have informed treatment decisions: patients with *KMT2B*-mediated early-onset dystonia were treated by deep brain stimulation[10]; individuals with *DIAPH1*-related macrothrombocytopenia and deafness[11] were treated for their thrombocytopenia in a preoperative setting with eltrombopag[12]; and a case of severe thrombocytopenia, myelofibrosis and bleeding owing to a gain-of-function mutation in *SRC*[13] was cured by an allogeneic haematopoietic stem cell transplant. Our diagnoses have stratified patient care: patients with primary immune disorders owing to variants in *NFKB1*, which we have shown are the most common monogenic cause of combined variable immunodeficiency[14], have unexplained splenomegaly and an increased risk of cancer; 27 cases with isolated thrombocytopenia caused by variants in *ANKRD26*, *ETV6* or *RUNX1* have an increased risk of malignancy[15–17] compared with 19 cases with thrombocytopenia caused by variants in *ACTN1*, *CYCS* or *TUBB1*. Our discoveries have also improved the accuracy of prognosis: we found that mutations in *BMPR2*[18] and *EIF2AK4*[19] carry a poorer-than-average prognosis in pulmonary arterial hypertension and we plan prognostication studies of four genes (*ATP13A3*, *AQP1*, *GDF2* and *SOX17*) that we recently reported are aetiological[20].

Genetic associations with rare diseases

Several cases with similar aetiologies are typically needed for discovery in rare disease genetics. Cases can be aggregated across distinct studies using Matchmaker Exchange[21]. We identified novel aetiologies for *SLC18A2*[22] and *WASF1*[23] using Matchmaker Exchange (Supplementary Information). However, in a study of a large unified health system, it is possible to make discoveries by statistical analyses of patient collections.

We applied BeviMed[24] to identify associations between genes and rare diseases under various modes of inheritance (Supplementary Information). We labelled groups of cases with a common tag if their phenotypes were a priori judged to be compatible with a shared aetiology (Supplementary Table 3). The number of unrelated cases with each tag ranged from three, for Roifman syndrome, to 1,101, for pulmonary arterial hypertension. We analysed each gene–tag pair independently and considered a posterior probability of association of greater than 0.75 to be strong evidence of a genetic aetiology. To account for correlation between tags, we recorded only one association per gene, corresponding to the tag for which the highest posterior probability of association was obtained. Conditional on gene causality for a tag, BeviMed reported posterior probabilities over the mode of inheritance, the molecular consequence class of variants that mediate disease risk (for example, variants in the 5′ untranslated region or predicted loss-of-function variants) and the pathogenicity of each specific variant.

We recorded strong evidence for association between 95 genes and 29 tags. The distribution of posterior probabilities implied a posterior estimate for the positive predictive value of 93%. The 95 genes included

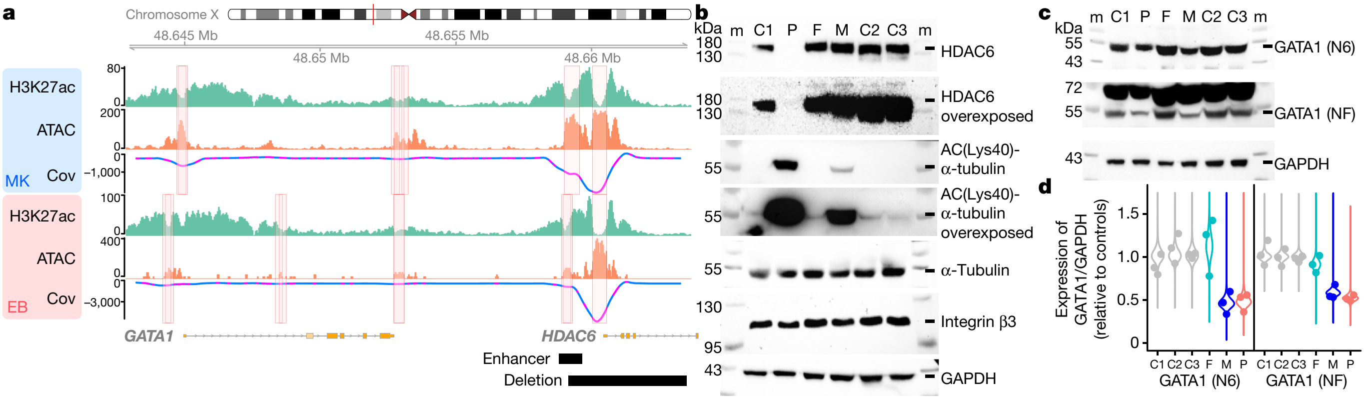


Fig. 4 | Causal variants in regulatory elements. a, From top to bottom: X chromosome ideogram; read coverage of H3K27ac ChIP–seq (green) and ATAC-seq (orange) in megakaryocytes (MK); the smoothed covariance (Cov) between H3K27ac ChIP–seq and ATAC-seq coverages for megakaryocytes, which were used to call regulatory elements (overlying coral rectangles); pink segments indicate regions in which the locally normalized ATAC-seq coverage exceeds the locally normalized H3K27ac ChIP–seq coverage (Supplementary Information); the corresponding three tracks and overlays for erythroblasts (EB); gene exons are shown in orange; the *GATA1* enhancer and the large deletion in the proband are shown as horizontal bars. A regulatory element overlapping the enhancer was identified by RedPop in megakaryocytes and erythroblasts but not in the other four cell types (tracks for these cell types are not shown). The deleted element binds to transcription factors that are characteristic of the megakaryocyte lineage: FLI1, GATA1/2, MEIS1, RUNX1 and TAL1 (binding not shown). **b**–**d**, P, proband; M, mother; F, father; C1, C2 and C3 are controls. **b**, **c**, m, marker. **b**, Representative immunoblots for total platelet lysates for the indicated proteins and individuals ($n = 2$). **c**, Representative example of $n = 3$ replicate immunoblots of total platelet lysates using two GATA1 antibodies (N6 and NF). **d**, Dot plots of GATA1 protein quantifications (as in **c**). The underlying violin plots show posterior predictive densities for the distribution of standardized GATA1 expression. The 90% credible intervals for the ratio of expression using the N6 antibody in father, mother, proband to the geometric mean in controls were 0.86–1.45, 0.35–0.59 and 0.37–0.62, respectively; similarly, for the expression using the NF antibody, the 95% credible intervals were 0.80–1.05, 0.51–0.67 and 0.45–0.60, respectively.

68 established DGGs, 11 DGGs that were discovered since 2015[10,14,20,25–30] and 16 candidates that require further investigation (Fig. 2d and Supplementary Table 3). Therefore, 79 of the 95 associations are confirmed, setting a lower bound on the true positive predictive value of 83%, which is broadly in line with an ancestry-controlled statistical estimate of the study-wide positive predictive value of 79% (Supplementary Information). We estimated that 611.3 cases can be explained by rare variants in the 79 confirmed genes, 115.6 of which are explained by the association between *BMPR2* and pulmonary arterial hypertension. Associations with 51 of the 95 genes relied solely on evidence from singleton variants, showing the power of joint statistical modelling of rare variants. Only three of the unconfirmed associations relied on evidence from alleles carried by more than one case, demonstrating the robustness of the results to cryptic relatedness. For one gene (*GP1BB*), the mode of inheritance inferred by BeviMed differed from that established in the literature, challenging long-held assumptions[31]. These results and other findings from this project[9–11,14,20,23,32–36], show that a unified analysis of homogeneously collected genetic and phenotypic data from a large phenotypically heterogeneous rare disease cohort is a powerful approach for genetic discovery.

Genetics of the tails of a quantitative trait

Several heritable rare diseases (for example, familial hypercholesterolaemia, combined variable immunodeficiency, thrombocytopenia and von Willebrand disease) are diagnosed and clinically characterized by reference to a quantitative trait that acts as a causal intermediate (or close proxy) for pathology. Alleles with large effects on a quantitative trait predispose carriers to lie in the extreme tails and hence to negative selection pressure. Consequently, such alleles are rare. We sought to identify genes that were likely to mediate red blood cell (RBC)-associated pathologies by WGS of UK Biobank participants in the tails of a univariate quantitative phenotype, computed to optimize rare variant heritability. We derived the univariate phenotype by considering the joint distribution of estimated effect sizes from GWAS associations between variants with minor allele frequencies of <1% and four RBC full blood count traits[7] (Fig. 3a). We sequenced 764 participants, 383 of whom were in the left tail of the phenotype, corresponding to a low RBC count and a high mean cell volume, and 381 of whom were in the right tail of the phenotype, corresponding to a high RBC count and a low mean cell volume (Fig. 3b, c).

The distribution of a polygenic predictor of the phenotype derived from an RBC full blood count GWAS exhibited left and right shifts from the population distribution in the respective tails (Fig. 3d). However, these shifts were less strong than predicted by Gaussian variance components modelling, a discrepancy that might be partly explained by rare alleles generating excess density in the tails (phenotype kurtosis = 6.9). A WGS GWAS of an ordinal outcome (left tail, unselected, right tail) did not yield novel associations. Therefore, we treated each of the tail groups as a set of cases in a BeviMed analysis and identified 12 genes with a posterior probability of association >0.4, which is a liberal threshold (Fig. 3e). *HBB* and *TFRC* can be considered causal, as known mutations cause microcytic anaemias. Other genes, including *CUX1* and *ALG1*, are plausible candidates. These results (Supplementary Table 3) indicate that the analysis of quantitative extremes in apparently healthy population samples may identify medically relevant loci[7,37].

Aetiological variants in regulatory elements

Rare variants in regulatory elements can cause disease by disrupting transcription or translation[38,39]. Recent studies have suggested that—at least in neurodevelopmental disorders—a small percentage of cases are attributable to de novo non-coding SNVs in regulatory elements that are active in relevant tissues[40]. Larger variants may be more disruptive to regulatory elements than SNVs. We searched for aetiological variants, including large deletions, in the regulatory elements of 246 DGGs implicated in recessive haematopoiesis-related disorders (Supplementary Information). First, we defined a set of active regulatory elements—a ‘regulome’—for each of six haematological cell types, by merging transcription-factor-binding sites identified by chromatin immunoprecipitation followed by sequencing (ChIP–seq) with genomic regions called by RedPop. RedPop is a detection method that uses the negative covariance between data from the assay for transposase-accessible chromatin using sequencing (ATAC-seq) and ChIP–seq coverage of histone H3 K27 acetylation (H3K27ac) in regulatory elements (Supplementary Information). We linked the regulatory elements to genes using genomic proximity and promoter capture chromosome conformation capture (pcHi-C)[41]. Second, we assigned each regulome to one or more of three rare disease domains—bleeding, thrombotic and platelet disorders, primary immune disorders and stem cell and myeloid disorders—according to the relevance of the corresponding cell types to the domains (Supplementary Table 3). Last, we searched for cases with a rare homozygous or hemizygous deletion of a regulatory element active in a relevant cell type and linked to a DGG of the domain of the case. We also searched for deletions that met these criteria that were in compound heterozygosity with a rare coding variant in a DGG linked to the deleted element. These approaches explained three cases: a patient with a primary immune disorder who carried a deletion overlapping the 5′ untranslated region of *ARPC1B* in compound heterozygosity with a frameshift variant in the same gene[36], a boy with autism spectrum disorder and thrombocytopenia who carried a hemizygous deletion of a *GATA1* enhancer and a patient with several autoimmune-mediated cytopenias who carried a homozygous deletion of an intronic CTCF-binding site[42] of *LRBA*.

The X-linked variant carried by the boy with autism spectrum disorder deleted a *GATA1* enhancer and exons 1–4 of *HDAC6* (Fig. 4 and Extended Data Fig. 9). He had a persistently low platelet count ($52 \times 10^9\ \mathrm{l}^{-1}$), an elevated mean platelet volume (15.1 fl) and normal RBC parameters except for mild dyserythropoiesis. Electron microscopy analyses showed lower than usual platelet α-granule content. Stem cell culture recapitulated poor platelet formation by megakaryocytes. These symptoms are typical of patients with a pathogenic coding *GATA1* allele[43]. His platelets contained abnormally low GATA1, consistent with weak transcription due to the deletion of the enhancer[44]. HDAC6 deacetylases Lys40 of α-tubulin, which localizes in polymerized microtubules[45]. The absence of HDAC6 was accompanied by an increase in acetylated α-tubulin in platelets. Knockout of the mouse homologue, *Hdac6*, causes aberrant acetylation of α-tubulin, which leads to bleeding[46] and abnormal behaviour[47]. Thus, the reduced expression of *GATA1* and the absence of HDAC6 jointly caused a previously undescribed syndrome of macrothrombocytopenia that is accompanied by neurodevelopmental problems. The patient with a homozygous deletion of a CTCF-binding site in the first intron of *LRBA* presented with autoantibody-mediated pancytopenia due to a loss of tolerance for multiple autoantigens, which is characteristic of impaired *LRBA* function[48].

We adapted our approaches for identifying pathogenic deletions in regulatory elements to identify pathogenic non-coding SNVs. We focused on SNVs with a combined annotation-dependent depletion (CADD)[49] score > 20 in compound heterozygosity with a high-impact coding variant in the assigned DGG. This approach identified two potentially aetiological SNVs in elements assigned to *AP3B1* and *MPL*. We studied the latter mutation (chromosome 1: 43803414G>A), carried by a 10-year-old boy, in more detail (Extended Data Fig. 10). *MPL* encodes the receptor for the megakaryocyte growth factor thrombopoietin[50]. Loss of *MPL* causes chronic amegakaryocytic thrombocytopenia[51]. The SNV was in a megakaryocyte-specific RedPop-identified regulatory element. It had CADD = 21.8, was absent from gnomAD and was in compound heterozygosity with a deletion of exon 10 of *MPL*. The mutant allele was associated with 50% reduced promoter activity, leading to a significant reduction in platelet MPL levels. In contrast to *MPL*-null patients[52], who are severely thrombocytopenic because their bone marrow is almost devoid of megakaryocytes, the patient had platelet counts of $45 \times 10^9\ \mathrm{l}^{-1}$ and a bone marrow that was only moderately depleted of megakaryocytes. As the regulatory SNV does not abolish *MPL* transcription completely, the boy has a milder clinical phenotype than *MPL*-null individuals.

Discussion

The resolution of unknown rare disease aetiologies will be hastened by the standardization and integration of clinical testing and research on a national scale. The NHS in England plans to increase provision of WGS-based diagnostics from 8,000 to 30,000 samples per month. To achieve this, it has reduced the number of clinical genomics laboratories to seven and introduced unified staff training in WGS, informatics and genomics. The development of statistical methodology to interpret the new data and participant consent to recall for follow-up experiments will be of critical importance. Additionally, long-read sequencing may be needed to overcome the difficulty of calling complex structural variants by WGS. We have initiated WGS of UK Biobank participants to identify rare variant associations with participants in the extreme tails of a quantitative phenotype who are typically excluded from GWAS. These associations can identify genes that mediate Mendelian pathologies. We have also shown that epigenetic data for cell types that mediate aetiology, combined with WGS, can identify regulatory elements that contain pathogenic non-coding mutations. The exploration of regulatory variation is a promising focus for future research and clinical intervention.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2434-2.

1. Ferreira, C. R. The burden of rare diseases. *Am. J. Med. Genet. A*. **179**, 885–892 (2019).
2. Bycroft, C. et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
3. Boycott, K. M. et al. International cooperation to enable the diagnosis of all rare genetic diseases. *Am. J. Hum. Genet*. **100**, 695–705 (2017).
4. Vissers, L. E. L. M. et al. A clinical utility study of exome sequencing versus conventional genetic testing in pediatric neurology. *Genet. Med*. **19**, 1055–1063 (2017).
5. Richards, S. et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med*. **17**, 405–423 (2015).
6. Van Houten, C. V. et al. Whole exome sequencing and characterization of coding variation in 49,960 individuals in the UK Biobank. Preprint at *bioRxiv* https://doi.org/10.1101/572347 (2019).
7. Astle, W. J. et al. The allelic landscape of human blood cell trait variation and links to common complex disease. *Cell* **167**, 1415–1429 (2016).
8. Belkadi, A. et al. Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. *Proc. Natl Acad. Sci. USA* **112**, 5473–5478 (2015).
9. Carss, K. J. et al. Comprehensive rare variant analysis via whole-genome sequencing to determine the molecular pathology of inherited retinal disease. *Am. J. Hum. Genet*. **100**, 75–90 (2017).
10. Meyer, E. et al. Mutations in the histone methyltransferase gene *KMT2B* cause complex early-onset dystonia. *Nat. Genet*. **49**, 223–237 (2017).
11. Stritt, S. et al. A gain-of-function variant in *DIAPH1* causes dominant macrothrombocytopenia and hearing loss. *Blood* **127**, 2903–2914 (2016).
12. Westbury, S. K. et al. Phenotype description and response to thrombopoietin receptor agonist in *DIAPH1*-related disorder. *Blood Adv*. **2**, 2341–2346 (2018).
13. Turro, E. et al. A dominant gain-of-function mutation in universal tyrosine kinase SRC causes thrombocytopenia, myelofibrosis, bleeding, and bone pathologies. *Sci. Transl. Med*. **8**, 328ra30 (2016).
14. Tuijnenburg, P. et al. Loss-of-function nuclear factor κB subunit 1 (*NFKB1*) variants are the most common monogenic cause of common variable immunodeficiency in Europeans. *J. Allergy Clin. Immunol*. **142**, 1285–1296 (2018).
15. Noris, P. et al. *ANKRD26*-related thrombocytopenia and myeloid malignancies. *Blood* **122**, 1987–1989 (2013).
16. Noetzli, L. et al. Germline mutations in *ETV6* are associated with thrombocytopenia, red cell macrocytosis and predisposition to lymphoblastic leukemia. *Nat. Genet*. **47**, 535–538 (2015).
17. Song, W. J. et al. Haploinsufficiency of *CBFA2* causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. *Nat. Genet*. **23**, 166–175 (1999).
18. Evans, J. D. et al. *BMPR2* mutations and survival in pulmonary arterial hypertension: an individual participant data meta-analysis. *Lancet Respir. Med*. **4**, 129–137 (2016).
19. Hadinnapola, C. et al. Phenotypic characterization of *EIF2AK4* mutation carriers in a large cohort of patients diagnosed clinically with pulmonary arterial hypertension. *Circulation* **136**, 2022–2033 (2017).

20. Gräf, S. et al. Identification of rare sequence variation underlying heritable pulmonary arterial hypertension. *Nat. Commun*. **9**, 1416 (2018).
21. Philippakis, A. A. et al. The Matchmaker Exchange: a platform for rare disease gene discovery. *Hum. Mutat*. **36**, 915–921 (2015).
22. Padmakumar, M. et al. A novel missense variant in *SLC18A2* causes recessive brain monoamine vesicular transport disease and absent serotonin in platelets. *JIMD Rep*. **47**, 9–16 (2019).
23. Ito, Y. et al. De novo truncating mutations in *WASF1* cause intellectual disability with seizures. *Am. J. Hum. Genet*. **103**, 144–153 (2018).
24. Greene, D., Richardson, S. & Turro, E. A fast association test for identifying pathogenic variants involved in rare diseases. *Am. J. Hum. Genet*. **101**, 104–114 (2017).
25. Merico, D. et al. Compound heterozygous mutations in the noncoding *RNU4ATAC* cause Roifman syndrome by disrupting minor intron splicing. *Nat. Commun*. **6**, 8718 (2015).
26. Ananth, A. L. et al. Clinical course of six children with *GNAO1* mutations causing a severe and distinctive movement disorder. *Pediatr. Neurol*. **59**, 81–84 (2016).
27. Horn, D. et al. Biallelic *COL3A1* mutations result in a clinical spectrum of specific structural brain anomalies and connective tissue abnormalities. Am. *J. Med. Genet. A*. **173**, 2534–2538 (2017).
28. Khan, S. Y. et al. Splice-site mutations identified in *PDE6A* responsible for retinitis pigmentosa in consanguineous Pakistani families. *Mol. Vis*. **21**, 871–882 (2015).
29. Petrovski, S. et al. Germline de novo mutations in GNB1 cause severe neurodevelopmental disability, hypotonia, and seizures. *Am. J. Hum. Genet*. **98**, 1001–1010 (2016).
30. Akawi, N. et al. Discovery of four recessive developmental disorders using probabilistic genotype and phenotype matching among 4,125 families. *Nat. Genet*. **47**, 1363–1369 (2015).
31. Sivapalaratnam, S. et al. Rare variants in *GP1BB* are responsible for autosomal dominant macrothrombocytopenia. *Blood* **129**, 520–524 (2017).
32. Westbury, S. K. et al. Expanded repertoire of *RASGRP2* variants responsible for platelet dysfunction and severe bleeding. *Blood* **130**, 1026–1030 (2017).
33. Pleines, I. et al. Mutations in tropomyosin 4 underlie a rare form of human macrothrombocytopenia. *J. Clin. Invest*. **127**, 814–829 (2017).
34. Heremans, J. et al. Abnormal differentiation of B cells and megakaryocytes in patients with Roifman syndrome. *J. Allergy Clin. Immunol*. **142**, 630–646 (2018).
35. Lentaigne, C. et al. Germline mutations in the transcription factor IKZF5 cause thrombocytopenia. *Blood* **134**, 2070–2081 (2019).
36. Thaventhiran, J. E. D. et al. Whole-genome sequencing of a sporadic primary immunodeficiency cohort. *Nature* (2020). https://doi.org/10.1038/s41586-020-2265-1
37. Natarajan, P. et al. Deep-coverage whole genome sequences and blood lipids among 16,324 individuals. *Nat. Commun*. **9**, 3391 (2018).
38. Giardine, B. et al. Updates of the HbVar database of human hemoglobin variants and thalassemia mutations. *Nucleic Acids Res*. **42**, D1063–D1069 (2014).
39. Albers, C. A. et al. Compound inheritance of a low-frequency regulatory SNP and a rare null mutation in exon-junction complex subunit *RBM8A* causes TAR syndrome. *Nat. Genet*. **44**, 435–439 (2012).
40. Short, P. J. et al. De novo mutations in regulatory elements in neurodevelopmental disorders. *Nature* **555**, 611–616 (2018).
41. Javierre, B. M. et al. Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters. *Cell* **167**, 1369–1384 (2016).
42. Ong, C. T. & Corces, V. G. CTCF: an architectural protein bridging genome topology and function. *Nat. Rev. Genet*. **15**, 234–246 (2014).
43. Freson, K. et al. Platelet characteristics in patients with X-linked macrothrombocytopenia because of a novel GATA1 mutation. *Blood* **98**, 85–92 (2001).
44. Fulco, C. P. et al. Systematic mapping of functional enhancer–promoter connections with CRISPR interference. *Science* **354**, 769–773 (2016).
45. Skultetyova, L. et al. Human histone deacetylase 6 shows strong preference for tubulin dimers over assembled microtubules. *Sci. Rep*. **7**, 11547 (2017).
46. Sadoul, K. et al. HDAC6 controls the kinetics of platelet activation. *Blood* **120**, 4215–4218 (2012).
47. Fukada, M. et al. Loss of deacetylation activity of Hdac6 affects emotional behavior in mice. *PLoS One* **7**, e30924 (2012).
48. Lopez-Herrera, G. et al. Deleterious mutations in *LRBA* are associated with a syndrome of immune deficiency and autoimmunity. *Am. J. Hum. Genet*. **90**, 986–1001 (2012).
49. Kircher, M. et al. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet*. **46**, 310–315 (2014).
50. Wendling, F. et al. cMpl ligand is a humoral regulator of megakaryocytopoiesis. *Nature* **369**, 571–574 (1994).
51. Tijssen, M. R. et al. Functional analysis of single amino-acid mutations in the thrombopoietin-receptor Mpl underlying congenital amegakaryocytic thrombocytopenia. *Br. J. Haematol*. **141**, 808–813 (2008).
52. Ballmaier, M. & Germeshausen, M. Congenital amegakaryocytic thrombocytopenia: clinical presentation, diagnosis, and treatment. *Semin. Thromb. Hemost*. **37**, 673–681 (2011).

---

[1]Department of Haematology, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK. [2]NIHR BioResource, Cambridge University Hospitals NHS Foundation, Cambridge Biomedical Campus, Cambridge, UK. [3]MRC Biostatistics Unit, Cambridge Institute of Public Health, University of Cambridge, Cambridge, UK. [4]NHS Blood and Transplant, Cambridge Biomedical Campus, Cambridge, UK. [5]Department of Medicine, School of Clinical Medicine, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK. [6]British Heart Foundation Cambridge Centre of Excellence, University of Cambridge, Cambridge, UK. [7]Department of Cardiovascular Sciences, Center for Molecular and Vascular Biology, KU Leuven, Leuven, Belgium. [8]Cambridge Institute of Therapeutic Immunology and Infectious Disease, Jeffrey Cheah Biomedical Centre, Cambridge Biomedical Campus, Cambridge, UK. [9]MRC Clinical Sciences Centre, Faculty of Medicine, Imperial College London, London, UK. [10]Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK. [11]The Nuffield Department of Clinical Neurosciences, University of Oxford, John Radcliffe Hospital, Oxford, UK. [12]NIHR Oxford Biomedical Research Centre, Oxford University Hospitals Trust, Oxford, UK. [13]High Performance Computing Service, University of Cambridge, Cambridge, UK. [14]Genomics England Ltd, London, UK. [15]William Harvey Research Institute, NIHR Biomedical Research Centre at Barts, Queen Mary University of London, London, UK. [16]Department of Clinical Neurosciences, School of Clinical Medicine, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK. [17]Medical Research Council Mitochondrial Biology Unit, Cambridge Biomedical Campus, Cambridge, UK. [18]Women and Children's Health, School of Life Course Sciences, King's College London, London, UK. [19]Department of Renal Medicine, University College London, London, UK. [20]Rare Renal Disease Registry, UK Renal Registry, Bristol, UK. [21]King's College London, London, UK. [22]Department of Paediatric Nephrology, Evelina London Children's Hospital, Guy's & St Thomas' NHS Foundation Trust, London, UK. [23]Department of Haematology, Hammersmith Hospital, Imperial College Healthcare NHS Trust, London, UK. [24]Centre for Haematology, Imperial College London, London, UK. [25]Department of Medical Genetics, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK. [26]NIHR Cambridge Biomedical Research Centre, Cambridge Biomedical Campus, Cambridge, UK. [27]Cancer Research UK Cambridge Centre, Cambridge Biomedical Campus, Cambridge, UK. [28]Stroke Research Group, Department of Clinical Neurosciences, University of Cambridge, Cambridge Biomedical Campus, Bristol, UK. [29]European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Cambridge, UK. [30]School of Cellular and Molecular Medicine, University of Bristol, Bristol, UK. [31]University Hospitals Bristol NHS Foundation Trust, Bristol, UK. [32]Department of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, UK. [33]MRC Molecular Haematology Unit, MRC Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, UK. [34]Department of Paediatrics, Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, UK. [35]Oxford University Hospitals NHS Foundation Trust, Oxford, UK. [36]Bristol Renal and Children's Renal Unit, Bristol Medical School, University of Bristol, Bristol, UK. [37]Bristol Royal Hospital for Children, University Hospitals Bristol NHS Foundation Trust, Bristol, UK. [38]Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK. [39]UCL Great Ormond Street Institute of Child Health, London, UK. [40]Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK. [41]Moorfields Eye Hospital NHS Trust, London, UK. [42]UCL Institute of Opthalmology, University College London, London, UK. [43]Department of Medicine, Imperial College London, London, UK. [44]Institute of Reproductive and Developmental Biology, Department of Surgery and Cancer, Faculty of Medicine, Hammersmith Hospital, Imperial College Healthcare NHS Trust, London, UK. [45]Illumina Cambridge, Little Chesterford, UK. [46]Addenbrookes Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK. [47]Department of Renal Medicine, Addenbrookes Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK. [48]Wellcome Sanger Institute, Cambridge, UK. *A list of authors and their affiliations appears in the online version of the paper. ✉e-mail: et341@cam.ac.uk; flr24@cam.ac.uk; who1000@cam.ac.uk

NIHR BioResource for the 100,000 Genomes Project

Stephen Abbs[49], Lara Abulhoul[50], Julian Adlard[51], Munaza Ahmed[52], Timothy J. Aitman[9,10], Hana Alachkar[53], David J. Allsup[54,55], Jeff Almeida-King[29], Philip Ancliff[50], Richard Antrobus[56], Ruth Armstrong[25,26,27], Gavin Arno[41,42], Sofie Ashford[2,38], William J. Astle[3,4], Anthony Attwood[1,2], Paul Aurora[50], Christian Babbs[12,33], Chiara Bacchelli[12,39], Tamam Bakchoul[57], Siddharth Banka[58,59], Tadbir Bariana[60,61], Julian Barwell[62,63], Joana Batista[1,2], Helen E. Baxendale[5,64,65,66], Phil L. Beales[50,67], David L. Bennett[11,12], David R. Bentley[45], Agnieszka Bierzynska[36], Tina Biss[68], Maria A. K. Bitner-Glindzicz[50,67], Graeme C. Black[58,59], Marta Bleda[5], Iulia Blesneac[11], Detlef Bockenhauer[50], Harm Bogaard[69], Christian J. Bourne[45], Sara Boyce[70], John R. Bradley[2,5,26,46,47], Eugene Bragin[71], Gerome Breen[72,73], Paul Brennan[74,75,76], Carole Brewer[77], Matthew Brown[1,2], Andrew C. Browning[78], Michael J. Browning[79], Rachel J. Buchan[80,81], Matthew S. Buckland[82], Teofila Bueser[83,84], Carmen Bugarin Diz[21], John Burn[76], Siobhan O. Burns[85,86], Oliver S. Burren[5,8], Nigel Burrows[46], Paul Calleja[13], Carolyn Campbell[87], Gerald Carr-White[88], Keren Carss[1,2], Ruth Casey[25,26,27], Mark J. Caulfield[14,15], Jenny Chambers[18,89], John Chambers[90,91,92,93,94], Melanie M. Y. Chan[19], Calvin Cheah[71], Floria Cheng[89], Patrick F. Chinnery[2,16,17], Manali Chitre[64], Martin T. Christian[95], Colin Church[96], Jill Clayton-Smith[58,59], Maureen Cleary[50], Naomi Clements Brod[1,2], Gerry Coghlan[82], Elizabeth Colby[36], Trevor R. P. Cole[97], Janine Collins[1,98], Peter W. Collins[99], Camilla Colombo[45], Cecilia J. Compton[83], Robin Condliffe[100], Stuart Cook[80,101,102,103], H. Terence Cook[104], Nichola Cooper[43], Paul A. Corris[105,106], Abigail Furnell[1,2], Fiona Cunningham[29], Nicola S. Curry[107], Antony J. Cutler[108], Matthew J. Daniels[32,35,110], Mehul Dattani[67,111], Louise C. Daugherty[1,2,14], John Davis[1,2], Anthony De Soyza[75,106,112], Sri V. V. Deevi[1,2], Timothy Dent[35], Charu Deshpande[83], Eleanor F. Dewhurst[1,2], Peter H. Dixon[18], Sofia Douzgou[58,59], Kate Downes[1,2], Anna M. Drazyk[28], Elizabeth Drewe[113], Daniel Duarte[1,2], Tina Dutt[114], J. David M. Edgar[115,116], Karen Edwards[1,2], William Egner[117], Melanie N. Ekani[88], Perry Elliott[118,119], Wendy N. Erber[120], Marie Erwood[1,2], Maria C. Estiu[121], Dafydd Gareth Evans[122], Gillian Evans[123], Tamara Everington[124,125], Mélanie Eyries[126,128], Hiva Fassihi[127], Remi Favier[129], Jack Findhammer[130], Debra Fletcher[1,2], Frances A. Flinter[83], R. Andres Floto[5,46,66], Tom Fowler[14,15], James Fox[1,2], Amy J. Frary[1,2], Courtney E. French[64], Kathleen Freson[7], Mattia Frontini[1,4,6], Daniel P. Gale[19,20], Henning Gall[131], Vijeya Ganesan[50], Michael Gattens[46], Claire Geoghegan[45], Terence S. A. Gerighty[45], Ali G. Gharavi[132], Stefano Ghio[133], Hossein-Ardeschir Ghofrani[43,131], J. Simon R. Gibbs[80], Kate Gibson[87], Kimberly C. Gilmour[39,50], Barbara Girerd[134,135,136], Nicholas S. Gleadall[1,2], Sarah Goddard[137], David B. Goldstein[138], Keith Gomez[60,61], Pavels Gordins[139], David Gosal[53], Stefan Gräf[1,2,5], Jodie Graham[140], Luigi Grassi[1,2], Daniel Greene[1], Lynn Greenhalgh[141], Andreas Greinacher[142], Paolo Gresele[143], Philip Griffiths[144,145], Sofia Grigoriadou[146], Russell J. Grocock[45], Detelina Grozeva[25], Mark Gurnell[5,46], Scott Hackett[147], Charaka Hadinnapola[5], William M. Hague[148], Rosie Hague[149], Matthias Haimel[1,2,5], Matthew Hall[113], Helen L. Hanson[150], Eshika Haque[83], Kirsty Harkness[151], Andrew R. Harper[32,40], Claire L. Harris[106,109], Daniel Hart[98], Ahamad Hassan[152], Grant Hayman[153], Alex Henderson[76], Archana Herwadkar[53], Jonathan Hoffman[97], Simon Holden[154], Rita Horvath[144,155], Henry Houlden[156], Arjan C. Houweling[223], Luke S. Howard[80,157], Fengyuan Hu[1,2], Gavin Hudson[145], Joseph Hughes[45], Aarnoud P. Huissoon[147], Marc Humbert[134,135,136], Sean Humphray[45], Sarah Hunter[2,45], Matthew Hurles[48], Melita Irving[83], Louise Izatt[83], Roger James[1,2], Sally A. Johnson[106,109,158], Stephen Jolles[159], Jennifer Jolley[1,2], Dragana Josifova[83], Neringa Jurkute[41,61], Tim Karten[130], Johannes Karten[130], Mary A. Kasanicki[46], Hanadi Kazkaz[160], Rashid Kazmi[70], Peter Kelleher[161,162], Anne M. Kelly[46], Wilf Kelsall[46], Carly Kempster[1,2], David G. Kiely[100], Nathalie Kingston[1,2], Robert Klima[13], Nils Koelling[163], Myrto Kostadima[1], Gabor Kovacs[164,165], Ania Koziell[21,22], Roman Kreuzhuber[1,2], Taco W. Kuijpers[166,167], Ajith Kumar[52], Dinakantha Kumararatne[168], Manju A. Kurian[169,170], Michael A. Laffan[23,24], Fiona Lalloo[59], Michele Lambert[171,172], Hana Lango Allen[1,2], Allan Lawrie[173], D. Mark Layton[23,24], Nick Lench[71], Claire Lentaigne[23,24], Tracy Lester[87], Adam P. Levine[19], Rachel Linger[2,38], Hilary Longhurst[174], Lorena E. Lorenzo[146], Eleni Louka[12,33], Paul A. Lyons[5,8], Rajiv D. Machado[175,176], Robert V. MacKenzie Ross[177], Bella Madan[178], Eamonn R. Maher[25,26,27], Jesmeen Maimaris[39], Samantha Malka[41,42], Sarah Mangles[125], Rutendo Mapeta[1,2], Kevin J. Marchbank[106,179], Stephen Marks[50], Hugh S. Markus[28], Hanns-Ulrich Marschall[180], Andrew Marshall[181,182,183], Jennifer Martin[2,5,38], Mary Mathias[184], Emma Matthews[185,186], Heather Maxwell[149], Paul McAlinden[106], Mark I. McCarthy[12,40,187], Harriet McKinney[1,2], Aoife McMahon[29], Stuart Meacham[1,2], Adam J. Mead[33], Ignacio Medina Castello[13], Karyn Megy[1,2], Sarju G. Mehta[154], Michel Michaelides[41,42], Carolyn Millar[23,24], Shehla N. Mohammed[83], Shahin Moledina[50], David Montani[134,135,136], Anthony T. Moore[41,42,188], Joannella Morales[29], Nicholas W. Morrell[2,5], Monika Mozere[19], Keith W. Muir[189], Andrew D. Mumford[30,31], Andrea H. Nemeth[11,190], William G. Newman[58,59], Michael Newnham[5,66], Sadia Noorani[191], Paquita Nurden[192], Jennifer O'Sullivan[178], Samya Obaji[193], Chris Odhams[14], Steven Okoli[12,33], Andrea Olschewski[164], Horst Olschewski[164,165], Kai Ren Ong[97], S. Helen Oram[194], Elizabeth Ormondroyd[12,32], Willem H. Ouwehand[1,2,4,6,48], Claire Palles[195], Sofia Papadia[2,38], Soo-Mi Park[26,27,49], David Parry[10], Smita Patel[196], Joan Paterson[25,26,27], Andrew Peacock[96], Simon H. Pearce[75,140], John Peden[45], Kathelijne Peerlinck[7], Christopher J. Penkett[1,2], Joanna Pepke-Zaba[66], Romina Petersen[1,2], Clarissa Pilkington[50], Kenneth E. S. Poole[5,46], Radhika Prathalingam[71], Bethan Psaila[12,33], Angela Pyle[145], Richard Quinton[75,145], Shamima Rahman[50,67], Stuart Rankin[13], Anupama Rao[50], F. Lucy Raymond[2,25], Paula J. Rayner-Matthews[1,2], Christine Rees[45], Augusto Rendon[1,14], Tara Renton[198], Christopher J. Rhodes[43], Andrew S. C. Rice[199,200], Sylvia Richardson[3], Alex Richter[56], Leema Robert[83], Irene Roberts[12,33,34], Anthony Rogers[71], Sarah J. Rose[83], Robert Ross-Russell[46], Catherine Roughley[123], Noemi B. A. Roy[12,33,35], Deborah M. Ruddy[83], Omid Sadeghi-Alavijeh[19], Moin A. Saleem[36,37], Nilesh Samani[202], Crina Samarghitean[1,2], Alba Sanchis-Juan[1,2], Ravishankar B. Sargur[117], Robert N. Sarkany[127], Simon Satchell[36,203], Sinisa Savic[204,205,206], John A. Sayer[75,145], Genevieve Sayer[83], Laura Scelsi[133], Andrew M. Schaefer[75,144], Sol Schulman[207], Richard Scott[14,50], Marie Scully[160], Claire Searle[208], Werner Seeger[131], Arjune Sen[12,209,210], W. A. Carrock Sewell[211], Denis Seyres[1,2], Neil Shah[39,50], Olga Shamardina[1,2], Susan E. Shapiro[107], Adam C. Shaw[83], Patrick J. Short[48], Keith Sibson[184], Lucy Side[212], Ilenia Simeoni[1,2], Michael A. Simpson[213], Matthew C. Sims[1,214], Suthesh Sivapalaratnam[4,98,215,216], Damian Smedley[14,15], Katherine R. Smith[14], Kenneth G. C. Smith[5,8], Katie Snape[150], Nicole Soranzo[1,48], Florent Soubrier[126], Laura Southgate[176,217], Olivera Spasic-Boskovic[49], Simon Staines[1,2], Emily Staples[5], Hannah Stark[2,38], Jonathan Stephens[1,2], Charles Steward[71], Kathleen E. Stirrups[1,2], Alex Stuckey[14], Jay Suntharalingam[177], Emilia M. Swietlik[5], Petros Syrris[118], R. Campbell Tait[218], Kate Talks[68], Rhea Y. Y. Tan[28], Katie Tate[71], John M. Taylor[87], Jenny C. Taylor[12,40], James E. Thaventhiran[8,219], Andreas C. Themistocleous[11], Ellen Thomas[14,15,88], David Thomas[5], Moira J. Thomas[220,221], Patrick Thomas[1,2], Kate Thomson[32,35], Adrian J. Thrasher[39], Glen Threadgold[29], Chantal Thys[7], Tobias Tilly[1,2], Marc Tischkowitz[25,26,49], Catherine Titterton[1,2], John A. Todd[108], Cheng-Hock Toh[114], Bas Tolhuis[130], Ian P. Tomlinson[195], Mark Toshner[5,66], Matthew Traylor[28], Carmen Treacy[5,66], Paul Treadaway[1,2], Richard Trembath[21], Salih Tuna[1,2], Wojciech Turek[13], Ernest Turro[1,2,3], Philip Twiss[49], Tom Vale[11], Chris Van Geet[7], Natalie van Zuydam[40,197], Maarten Vandekuilen[130], Anthony M. Vandersteen[222], Marta Vazquez-Lopez[89], Julie von Ziegenweidt[1,2], Anton Vonk Noordegraaf[69], Annette Wagner[46], Quinten Waisfisz[223], Suellen M. Walker[39,50], Neil Walker[1,2], Klaudia Walter[48], James S. Ware[80,81,101], Hugh Watkins[32,35,40], Christopher Watt[1,2], Andrew R. Webster[41,42], Lucy Wedderburn[39,224,225], Wei Wei[16,17], Steven B. Welch[226], Julie Wessels[137], Sarah K. Westbury[30,31], John-Paul Westwood[160], John Wharton[43], Deborah Whitehorn[1,2], James Whitworth[25,26,27], Andrew O. M. Wilkie[163], Martin R. Wilkins[43], Catherine Williamson[18,44], Brian T. Wilson[52,75,145], Edwin K. S. Wong[75,106], Nicholas Wood[156,227], Yvette Wood[1,2], Christopher Geoffrey Woods[25,46], Emma R. Woodward[59], Stephen J. Wort[81,228], Austen Worth[50], Michael Wright[76], Katherine Yates[1,2,5], Patrick F. K. Yong[229], Timothy Young[1,2], Ping Yu[1,2], Patrick Yu-Wai-Man[16,17,230] & Eliska Zlamalova[1]

[49]East Anglian Medical Genetics Service, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK. [50]Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK. [51]Yorkshire Regional Genetics Service, Chapel Allerton Hospital, Leeds Teaching Hospitals NHS Trust, Leeds, UK. [52]North East Thames Regional Genetics Service, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK. [53]Salford Royal NHS Foundation Trust, Salford, UK. [54]Queens Centre for Haematology and Oncology, Castle Hill Hospital, Hull and East Yorkshire NHS Trust, Cottingham, UK. [55]Hull York Medical School, University of Hull, Hull, UK. [56]University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK. [57]Center for Clinical Transfusion Medicine, University Hospital of Tübingen, Tubingen, Germany. [58]Evolution and Genomic Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK. [59]Manchester Centre for Genomic Medicine, St Mary's Hospital, Manchester Universities Foundation NHS Trust, Manchester, UK. [60]The Katharine Dormandy Haemophilia Centre and Thrombosis Unit, Royal Free London NHS Foundation Trust, London, UK. [61]University College London, London, UK. [62]Department of Clinical Genetics, Leicester Royal Infirmary, University Hospitals of Leicester, Leicester, UK. [63]University of Leicester, Leicester, UK. [64]Department of Paediatrics, School of Clinical Medicine, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK. [65]Division of Clinical Biochemistry and Immunology, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK. [66]Royal Papworth Hospital NHS Foundation Trust, Cambridge, UK. [67]Genetics and Genomic Medicine Programme, UCL Great Ormond Street Institute of Child Health, London, UK. [68]Haematology Department, Royal Victoria Infirmary, The Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK. [69]Department of Pulmonary Medicine, Amesterdam University Medical Centres, VU University Medical Centre, Amsterdam, The Netherlands. [70]Southampton General Hospital, University Hospital Southampton NHS Foundation Trust, Southampton, UK. [71]Congenica, Biodata Innovation Centre, Cambridge, UK. [72]MRC Social, Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK. [73]NIHR Biomedical Research Centre for Mental Health, Maudsley Hospital, London, UK. [74]Newcastle University, Newcastle upon Tyne, UK. [75]Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK. [76]Northern Genetics Service, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK. [77]Department of Clinical Genetics, Royal Devon & Exeter Hospital, Royal Devon and Exeter NHS Foundation Trust, Exeter, UK. [78]Newcastle Eye Centre, Royal Victoria Infirmary, The Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK. [79]Department of Immunology, Leicester Royal Infirmary, Leicester, UK. [80]National Heart and Lung Institute, Imperial College London, London, UK. [81]Royal Brompton Hospital, Royal Brompton and Harefield NHS Foundation Trust, London, UK. [82]Royal Free London NHS Foundation Trust, London, UK. [83]Clinical Genetics Department, Guy's and St Thomas NHS Foundation Trust, London, UK. [84]Florence Nightingale Faculty of Nursing, Midwifery & Palliative Care, King's College London, London, UK. [85]Institute of Immunity and Transplantation, University College London, London, UK. [86]Department of Immunology, Royal Free London NHS Foundation Trust, London, UK. [87]Oxford Medical Genetics Laboratories, Oxford University Hospitals NHS Foundation Trust, Oxford, UK. [88]Guy's and St Thomas' Hospital, Guy's and St Thomas' NHS Foundation Trust, London, UK. [89]Women's Health Research Centre, Department of Surgery and Cancer, Faculty of Medicine, Hammersmith Hospital, Imperial College Healthcare NHS Trust, London, UK. [90]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore. [91]Department of Epidemiology and Biostatistics, Imperial College London, London, UK. [92]Department of Cardiology, Ealing Hospital, London, UK. [93]Imperial College Healthcare NHS Trust, London, UK. [94]MRC-PHE Centre for Environment and Health, Imperial College London, London, UK. [95]Children's Renal and Urology Unit, Nottingham Children's Hospital, QMC, Nottingham University Hospitals NHS Trust, Nottingham, UK. [96]Golden Jubilee

National Hospital, Glasgow, UK. [97]West Midlands Regional Genetics Service, Birmingham Women's and Children's NHS Foundation Trust, Birmingham, UK. [98]The Royal London Hospital, Barts Health NHS Foundation Trust, London, UK. [99]Institute of Infection and Immunity, School of Medicine Cardiff University, Cardiff, UK. [100]Sheffield Pulmonary Vascular Disease Unit, Royal Hallamshire Hospital NHS Foundation Trust, Sheffield, UK. [101]MRC London Institute of Medical Sciences, Imperial College London, London, UK. [102]National Heart Research Institute Singapore, National Heart Centre Singapore, Singapore, Singapore. [103]Division of Cardiovascular and Metabolic Disorders, Duke-National University of Singapore, Singapore, Singapore. [104]Department of Immunology and Inflammation, Faculty of Medicine, Imperial College London, London, UK. [105]National Pulmonary Hypertension Service (Newcastle), The Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK. [106]Translational and Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK. [107]Oxford Haemophilia and Thrombosis Centre, Oxford University Hospitals NHS Trust, Oxford Comprehensive Biomedical Research Centre, Oxford, UK. [108]JDRF/Wellcome Diabetes and Inflammation Laboratory, Wellcome Centre for Human Genetics, Nuffield Department of Medicine, NIHR Oxford Biomedical Research Centre, University of Oxford, Oxford, UK. [109]The National Renal Complement Therapeutics Centre, Royal Victoria Infirmary, Newcastle upon Tyne, UK. [110]Department of Biotechnology, Graduate School of Engineering, Osaka University, Suita, Osaka, Japan. [111]London Centre for Paediatric Endocrinology and Diabetes, Great Ormond Street Hospital for Children, London, UK. [112]NIHR Centre for Aging, Newcastle University, Newcastle upon Tyne, UK. [113]Nottingham University Hospitals NHS Trust, Nottingham, UK. [114]The Roald Dahl Haemostasis and Thrombosis Centre, The Royal Liverpool University Hospital, Liverpool, UK. [115]St James's Hospital, Dublin, Ireland. [116]Trinity College Dublin, Dublin, Ireland. [117]Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK. [118]UCL Institute of Cardiovascular Science, University College London, London, UK. [119]Barts Heart Centre, St Bartholomew's Hospital, Barts Health NHS Trust, London, UK. [120]Medical School and School of Biomedical Sciences, Faculty of Health and Medical Sciences, The University of Western Australia, and PathWest Laboratory Medicine, Crawley, Western Australia, Australia. [121]Ramón Sardá Mother's and Children's Hospital, Buenos Aires, Argentina. [122]Manchester University NHS Foundation Trust, Manchester, UK. [123]Haemophilia Centre, Kent & Canterbury Hospital, East Kent Hospitals University Foundation Trust, Canterbury, UK. [124]Salisbury District Hospital, Salisbury NHS Foundation Trust, Salisbury, UK. [125]Haemophilia, Haemostasis and Thrombosis Centre, Hampshire Hospitals NHS Foundation Trust, Basingstoke, UK. [126]Departement de Genetique & ICAN, Hopital Pitie-Salpetriere, Assistance Publique Hopitaux de Paris, Paris, France. [127]St Johns Institute of Dermatology, Guy's and St Thomas' NHS Foundation Trust, London, UK. [128]UMRS 1166-ICAN, INSERM, UPMC, Sorbonne Universités, Paris, France. [129]Service d'Hematologie biologique, Centre de Reference des Pathologies Plaquettaires, Hopital Armand Trousseau, Assistance Publique-Hopitaux de Paris, Paris, France. [130]GENALICE, Harderwijk, The Netherlands. [131]University of Giessen and Marburg Lung Center (UGMLC), Giessen, Germany. [132]Division of Nephrology and Center for Precision Medicine and Genomics, Department of Medicine Columbia University Vagelos College of Physicians and Surgeons, New York, NY, USA. [133]Division of Cardiology, Fondazione IRCCS Policlinico S. Matteo, Pavia, Italy. [134]Université Paris-Sud, Faculty of Medicine, University Paris-Saclay, Le Kremlin Bicetre, France. [135]Service de Pneumologie, Centre de Reference de l'Hypertension Pulmonaire, Hopital Bicetre (Assistance Publique Hopitaux de Paris), Le Kremlin Bicetre, France. [136]INSERM U999, Hospital Marie Lannelongue, Le Plessis Robinson, France. [137]University Hospitals of North Midlands NHS Trust, Stoke-on-Trent, UK. [138]Institute of Genomic Medicine and the Department of Genetics and Development, Columbia University Vagelos College of Physicians and Surgeons, New York, NY, USA. [139]East Yorkshire Regional Adult Immunology and Allergy Unit, Hull Royal Infirmary, Hull and East Yorkshire Hospitals NHS Trust, Hull, UK. [140]Newcastle BRC, Newcastle University, Newcastle upon Tyne, UK. [141]Department of Clinical Genetics, Liverpool Women's NHS Foundation, Liverpool, UK. [142]Institute for Immunology and Transfusion Medicine, University Medicine Greifswald, Greifswald, Germany. [143]Section of Internal and Cardiovascular Medicine, University of Perugia, Perugia, Italy. [144]Wellcome Centre for Mitochondrial Research, Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK. [145]Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK. [146]Barts Health NHS Foundation Trust, London, UK. [147]Birmingham Heartlands Hospital, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK. [148]Robinson Research Institute, Discipline of Obstetrics and Gynaecology, The University of Adelaide, Women's and Children's Hospital, Adelaide, South Australia, Australia. [149]Royal Hospital for Children, NHS Greater Glasgow and Clyde, Glasgow, UK. [150]Department of Clinical Genetics, St George's University Hospitals NHS Foundation Trust, London, UK. [151]Department of Neurology, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK. [152]Department of Neurology, Leeds Teaching Hospital NHS Trust, Leeds, UK. [153]Epsom & St Helier University Hospitals NHS Trust, London, UK. [154]Department of Clinical Genetics, Addenbrookes Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK. [155]John Walton Muscular Dystrophy Research Centre, Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK. [156]Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK. [157]National Pulmonary Hypertension Service, Imperial College Healthcare NHS Trust, London, UK. [158]Department of Paediatric Nephrology, Great North Children's Hospital, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK. [159]Immunodeficiency Centre for Wales, University Hospital of Wales, Cardiff, UK. [160]University College London Hospitals NHS Foundation Trust, London, UK. [161]Centre for Immunology & Vaccinology, Department of Medicine, Chelsea & Westminster Hospital, Imperial College London, London, UK. [162]Department of Respiratory Medicine, Royal Brompton & Harefield NHS Foundation Trust, London, UK. [163]MRC Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford, UK. [164]Ludwig Boltzmann Institute for Lung Vascular Research, Graz, Austria. [165]Department of Internal Medicine, Division of Pulmonology, Medical University of Graz, Graz, Austria. [166]Department of Pediatric Hematology, Immunology, Rheumatology and Infectious Diseases, Emma Children's Hospital, Academic Medical Center (AMC), University of Amsterdam, Amsterdam, The Netherlands. [167]Department of Blood Cell Research, Sanquin, Amsterdam, The Netherlands. [168]Department of Clinical Immunology, Addenbrookes Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK. [169]Developmental Neurosciences, UCL Great Ormond Street Institute of Child Health, London, UK. [170]Department of Neurology, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK. [171]Division of Hematology, The Children's Hospital of Philadelphia, Philadelphia, PA, USA. [172]Department of Pediatrics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA. [173]Department of Infection, Immunity & Cardiovascular Disease, University of Sheffield, Sheffield, UK. [174]Department of Specialist Allergy and Clinical Immunology, University College Hospital, University College London Hospitals NHS Foundation Trust, London, UK. [175]School of Life Sciences, University of Lincoln, Lincoln, UK. [176]Molecular and Clinical Sciences Research Institute, St George's University of London, London, UK. [177]Royal United Hospitals Bath NHS Foundation Trust, Bath, UK. [178]Department of Haematology, Guy's and St Thomas' NHS Foundation Trust, London, UK. [179]The National Renal Complement Therapeutics Centre, Royal Victoria Infirmary, Newcastle upon Tyne, UK. [180]Department of Molecular and Clinical Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden. [181]Faculty of Biology, Medicine and Health, School of Biological Sciences, Division of Neuroscience and Experimental Psychology, University of Manchester, Manchester, UK. [182]Department of Clinical Neurophysiology, Manchester University NHS Foundation Trust, Manchester, UK. [183]National Institute for Health Research/Wellcome Trust Clinical Research Facility, Manchester, UK. [184]Department of Haematology, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK. [185]The National Hospital for Neurology and Neurosurgery, University College London Hospitals NHS Foundation Trust, London, UK. [186]MRC Centre for Neuromuscular Diseases, Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK. [187]Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Churchill Hospital, Oxford University Hospitals NHS Trust, Oxford, UK. [188]Ophthalmology Department, UCSF School of Medicine, San Francisco, CA, USA. [189]Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, UK. [190]Department of Clinical Genetics, Churchill Hospital, Oxford University Hospitals NHS Trust, Oxford, UK. [191]Sandwell and West Birmingham Hospitals NHS Trust, Birmingham, UK. [192]Institut Hospitalo-Universitaire de Rythmologie et de Modelisation Cardiaque, Plateforme Technologique d'Innovation Biomedicale, Hopital Xavier Arnozan, Pessac, France. [193]The Arthur Bloom Haemophilia Centre, University Hospital of Wales, Cardiff, UK. [194]Department of Paediatric Haematology, University Hospital Southampton NHS Foundation Trust, Southampton, UK. [195]Institute of Cancer and Genomic Sciences, Institute of Biomedical Research, University of Birmingham, Birmingham, UK. [196]Department of Clinical Immunology, John Radcliffe Hospital, Oxford University Hospitals NHS Foundation Trust, Oxford, UK. [197]Oxford Centre for Diabetes, Endocrinology and Metabolism, Radcliffe Department of Medicine, University of Oxford, Oxford, UK. [198]King's College Hospital NHS Foundation Trust, London, UK. [199]Pain Research, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London, UK. [200]Pain Medicine, Chelsea and Westminster Hospital NHS Foundation Trust, London, UK. [201]Department of Haematology, Oxford University Hospital Foundation Trust, Oxford, UK. [202]Department of Cardiovascular Sciences and NIHR Leicester Biomedical Research Centre, University of Leicester, Leicester, UK. [203]North Bristol NHS Trust, Bristol, UK. [204]Department of Clinical Immunology and Allergy, St James's University Hospital, Leeds, UK. [205]The NIHR Leeds Biomedical Research Centre, Leeds, UK. [206]Leeds Institute of Rheumatic and Musculoskeletal Medicine, Leeds, UK. [207]Beth Israel Deaconess Medical Centre and Harvard Medical School, Boston, MA, USA. [208]Department of Clinical Genetics, Nottingham University Hospitals NHS Trust, Nottingham, UK. [209]Oxford Epilepsy Research Group, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK. [210]Department of Neurology, John Radcliffe Hospital, Oxford, UK. [211]Scunthorpe General Hospital, Northern Lincolnshire and Goole NHS Foundation Trust, Scunthorpe, UK. [212]Wessex Clinical Genetics Service, University Hospital Southampton NHS Foundation Trust, Southampton, UK. [213]Genetics and Molecular Medicine, King's College London, London, UK. [214]Oxford Haemophilia and Thrombosis Centre, Churchill Hospital, Oxford University Hospitals NHS Trust, Oxford, UK. [215]Department of Haematology, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK. [216]Queen Mary University of London, London, UK. [217]Faculty of Life Sciences and Medicine, King's College London, London, UK. [218]Glasgow Royal Infirmary, NHS Greater Glasgow and Clyde, Glasgow, UK. [219]MRC Toxicology Unit, School of Biological Sciences, University of Cambridge, Cambridge, UK. [220]Gartnavel General Hospital, NHS Greater Glasgow and Clyde, Glasgow, UK. [221]Queen Elizabeth University Hospital, Glasgow, UK. [222]Division of Medical Genetics, IWK Health Centre, Dalhousie University, Halifax, Nova Scotia, Canada. [223]Department of Clinical Genetics, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands. [224]NIHR Great Ormond Street Biomedical Research Centre, London, UK. [225]Arthritis Research UK Centre for Adolescent Rheumatology, University College London, London, UK. [226]Birmingham Chest Clinic and Heartlands Hospital, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK. [227]UCL Genetics Institute, UCL Division of Biosciences, University College London, London, UK. [228]Imperial College London, London, UK. [229]Frimley Park Hospital, NHS Frimley Health Foundation Trust, Camberley, UK. [230]NIHR Biomedical Research Centre at Moorfields Eye Hospital, UCL Institute of Ophthalmology, London, UK.
*A full list of members and their affiliations appears in the Supplementary Information.

Methods

Enrolment, research ethics and consent

Study participants were enrolled by one of three mechanisms between December 2012 and March 2017 under the overall coordination of the National Institute for Health Research BioResource (NBR) at Cambridge University Hospitals. Patients with rare diseases and their close relatives were enrolled into 15 rare disease domains approved by the Sequencing and Informatics Committee of the NBR. Enrolment of controls was coordinated by the University of Cambridge. Enrolment in the GEL domain was coordinated by Genomics England Ltd. Enrolment in the UK Biobank (extreme red blood cell traits) (UKB) domain was jointly coordinated by the NBR and UK Biobank[2]. Participants in the rare disease domains were recruited mainly at NHS Hospitals in the United Kingdom, but also at hospitals overseas (Extended Data Fig. 1a and Supplementary Table 1). All 13,187 participants provided written informed consent, either under the East of England Cambridge South national research ethics committee (REC) reference no. 13/EE/0325 or under ethics for other REC-approved studies. Obtaining consent for overseas samples was the responsibility of the respective principal investigators at the hospitals at which enrolment took place. The NBR retained blank versions of the consent forms from overseas participants and a material transfer agreement was applied to regulate the exchange of samples and data between the donor institutions and the University of Cambridge.

Clinical and laboratory phenotype data

Staff at hospitals responsible for enrolment were provided with the eligibility criteria for their respective domains as described in the domain descriptions (Supplementary Information). The clinical and laboratory phenotype data were captured through case report forms by paper questionnaires or by online data capture applications and deposited in the NBR study database. Online data capture allowed for the free entry of HPO terms[53] by staff at the enrolment centre and data from paper questionnaires were transformed into HPO terms by the study coordination office. Free text entries were transformed into HPO terms where feasible. An overview of the HPO data obtained for the NBR rare disease domains is depicted in Extended Data Fig. 1c.

DNA sequencing

Pre-extracted DNA samples or EDTA-treated whole-blood samples were delivered to the NBR laboratory at Cambridge, where DNA was extracted from the whole blood. Samples were tested for adequate concentration (Picogreen), quality controlled for DNA degradation (gel electrophoresis) and purity (using ratio of the absorbance at 260 and 280 nm ($A_{260/280}$); Trinean) before selection for WGS. DNA samples were prepared at a minimum concentration of 30 ng μl^{-1} in 110 µl, visually inspected for degradation and had to have an $A_{260/280}$ between 1.75 and 2.04. They were then prepared in batches of 96 and shipped on dry ice to the sequencing provider (Illumina). Further sample quality control was performed by Illumina to ensure that the concentration of the DNA was >30 ng μl^{-1} and that every sample generated high-quality microarray genotyping data (Illumina Infinium Human Core Exome microarray). Samples with a repeated array genotyping call rate <0.99, high levels of cross-contamination, mismatches with the declared gender that could not be resolved by further investigation, or for which consent had been withdrawn, were excluded from WGS ($n = 59$). The genotyping data were also used for positive sample identification before data delivery. For each sample, 0.5 µg of DNA was fragmented using Covaris LE220 (Covaris) to obtain an average size of 450-bp DNA fragments. DNA samples were processed using the Illumina TruSeq DNA PCR-Free Sample Preparation kit (Illumina) on the Hamilton Microlab Star (Hamilton Robotics). The final libraries were checked using the Roche LightCycler 480 II (Roche Diagnostics) with KAPA Library Quantification Kit (Kapa Biosystems) for concentration. From February 2014 to June 2017, three read lengths were used: 100 bp, 125 bp and 150 bp (377, 3,154 and 9,656 samples, respectively). Samples sequenced with 100-bp and 125-bp reads used three and two lanes of an Illumina HiSeq 2500 instrument, respectively, while samples sequenced with 150-bp reads used a single lane of a HiSeq X instrument. At least 95% of the autosomal genome had to be covered at 15× and a maximum of 5% of insert sizes had to be less than twice the read length. Following sample and data quality control at Illumina, 13,187 sets of WGS data files were received by the University of Cambridge High Performance Computing Service (HPC) for further quality control.

WGS data-processing pipeline

The WGS data for the 13,187 samples returned by the sequencing provider underwent a series of processing steps (Extended Data Fig. 2), described in detail in the Supplementary Information. In brief, the samples were sex karyotyped and pairwise kinship coefficients were computed. This information was used to check for repeat sample submissions and sample swaps. Additionally, four further quality control checks were applied to ensure the SNV and indel call data were of a high standard. Overall, 150 samples (1.1%) were removed, leaving a dataset of 13,037 samples for downstream analysis. The 13,037 individuals were assigned to one of the following ethnicities: 'European', 'African', 'South Asian', 'East Asian' or 'other'. Pairwise relatedness adjusted for population stratification was then computed and used to generate networks of closely related individuals and to define an MSUP of 10,259 individuals. The variants in the 13,037 individuals were left-aligned and normalized with bcftools, loaded into our HBase database and filtered on their overall pass rate, as defined in the Supplementary Information. The sex karyotypes, the ethnicities and the relatedness estimates were used, along with enrolment information, to annotate the samples and variants. Samples were annotated with: affected or unaffected status, membership of the set of probands, membership of the MSUP, ethnicity and sex karyotype. Variants were annotated with consequence predictions, HGMD information (where available) and population-specific allele frequencies.

Pertinent findings

For each of the 15 rare disease domains (that is, all domains except UKB, GEL and a domain comprising technical controls) a list of DGGs was generated by domain-specific experts. Genes were included in the lists if there was a high enough level of evidence in the literature for gene–disease association. The 2,497 gene–domain pairs, encompassing 2,073 unique DGGs across all domains, were manually curated and annotated with the relevant RefSeq and/or Ensembl transcript identifiers to support variant reporting. Transcripts were selected on the basis of (by order of priority) community input, presence in the Locus Reference Genomic resource[54] or designation as 'canonical' in Ensembl. Variants (SNVs, indels) were shortlisted if (1) their MAFs in control populations[55] were less than 1 in 1,000 for putative novel causal variants and less than 25 in 1,000 for variants listed as disease-causing in HGMD; (2) their predicted impact according to the Variant Effect Predictor[56] was 'high' or 'moderate' or if the consequences with respect to the designated transcript included one of 'splice_region_variant' or 'non_coding_transcript_exon_variant' if the variant was in a non-coding gene; (3) the variant affected a DGG relevant to the disease of the patient. Variants with more than three alleles or a MAF ≥10% in the cohort were discarded, respectively, to guard against errors in repetitive regions and to remove potential systematic artefacts. The above filtering criteria were applied universally to all domains except for ICP, which adopted a higher MAF threshold of 3% for both novel and previously reported variants. The higher threshold prevented erroneous filtering of causal variants carried at elevated frequencies by the male and non-child bearing female population. This strategy reduced the number of variants for review by the MDTs from about 4 million per person to fewer than 10 per person, while retaining almost all known regulatory or moderately common pathogenic variants. For each affected participant with prioritized variants, the variant calls, HPO-coded phenotype and the relevant metadata (unique study numbers; referring clinician and hospital; self-declared gender and genetically inferred sex, ancestry, relatedness, and consanguinity level) were transferred to Congenica for visualization in the Sapientia web application during MDT meetings. MDTs comprised experts from different hospitals across the United Kingdom and abroad, and typically consisted of an experienced clinician with domain-specific knowledge, a scientist with experience in clinical genomics, a clinical bioinformatician and a member of the reporting team. Assignment of the level of pathogenicity followed the American College of Medical Genetics guidelines[5] and variants were marked in Sapientia as pathogenic, likely pathogenic or of uncertain significance. Only pathogenic and likely pathogenic variants were systematically reported and variants of uncertain significance were reported at the discretion of the MDT. As per the REC-approved study protocol, secondary findings (for example, breast cancer pathogenic variants in *BRCA1* in patients not presenting with this phenotype) were not reported.

Genetic association testing in genes

We used the BeviMed statistical method[24] to identify genetic associations with rare diseases in our dataset. Each run of BeviMed requires the definition of a set of cases and controls, all of which should be unrelated to each other, and a set of rare variants to include in the inference. To achieve adequate power, the cases should be chosen such that they potentially share a common genetic aetiology (for example, because the phenotypes are similar) and the rare variants should be chosen such that they potentially share a mechanism of action on the phenotype (for example, because they are predicted to have a similar effect on a particular gene product). BeviMed computes posterior probabilities of no association, dominant association and recessive association and, conditional on dominant or recessive association, it computes the posterior probability that each variant is pathogenic. We can impose a prior correlation structure on the pathogenicity of the variants that reflects competing hypotheses as to which class of variant is responsible for disease. These classifications typically group variants by their predicted consequences. The class of variant responsible can then be inferred by BeviMed, thereby suggesting a particular aetiological mechanism. The BeviMed computed posterior probabilities can be used to estimate the number of cases attributable to variants in each gene, conditional on gene causality. The methodology is described in further detail in the Supplementary Information and in the original BeviMed publication[24]. BeviMed was applied gene-wise to infer associations between the genotypes of filtered rare variants and various case–control groupings (tags). For a given gene, only the maximum posterior probability over tags was recorded, to account for correlation between tags.

Regulome analysis

We applied the BLUEPRINT protocol for ChIP–seq data analysis (http://dcc.blueprint-epigenome.eu/#/md/chip_seq_grch37). We defined regulomes for activated CD4$^+$ T cells, B cells, erythroblasts, megakaryocytes, monocytes and resting CD4$^+$ T cells. For each cell type, we used open chromatin data (ATAC-seq or DNase-seq) and histone-modification data (H3K27ac) to identify regulatory elements using the RedPop method (Supplementary Information). Additionally, for megakaryocytes and erythroblasts, we had access to the following transcription-factor ChIP–seq data, which were used to call peaks and supplement the regulomes: FLI1, GATA1, GATA2, MEIS1, RUNX1, TAL1 and CTCF for megakaryocytes; GATA1, KLF1, NFE2 and TAL1 for erythroblasts; and CTCF for monocytes and B cells. For each cell type, the regulome build process proceeded as follows: (1) call RedPop regions using ATAC-seq or DNase-seq and H3K27ac-seq data; (2) call transcription factor and CTCF-binding peaks using ChIP–seq data if available and obtain enrichment scores; (3) discard peaks with an enrichment score <10 unless they overlap at least two other peaks; (4) collapse overlapping features to obtain a single genomic track; (5) merge features within 100 bp of each other. Each regulome feature was assigned a gene label using either gene annotations from Ensembl (v.75) or a compendium of previously published pcHi-C[41] as follows: (1) assign to a gene if the feature overlaps the gene or the region up to 10 kb either side of the gene body; (2) assign to a gene if the feature overlaps the pcHi-C 'blind' spot of the gene (this region is defined by three HindIII restriction fragments, incorporating the capture fragment overlapping the transcription start site of the target gene, and the 5′ and 3′ adjacent fragments); (3) assign to a gene if the feature overlaps a linked promoter-interacting region identified using pcHi-C in the same cell type.

Functional analysis of the *GATA1* enhancer and *HDAC6* deletion

The *GATA1* enhancer and *HDAC6* deletion was confirmed by PCR using primers HDAC6-F: 5′-CATCTTCAAGAGGATCAGAGG-3′ and HDAC6-R: 5′-CATAGCTAGACACTGGTT-3′. Electron microscopy analysis of platelets was performed as described previously[43]. Immunostaining of resting and fibrinogen spread platelets was performed as described previously[34] and analysed by structured illumination microscopy (SIM, Elyra S.1, Zeiss). Total protein lysates were obtained from platelets for immunoblot analysis as described previously[57]. The following antibodies were used for SIM and immunoblot analysis: rabbit anti-HDAC6 (clone D2E5, Cell Signaling Technology), mouse anti-acetylated tubulin antibody (clone 6-11B-1, Sigma), mouse anti-α-tubulin (A11126, Thermo Fisher Scientific), rabbit anti-VWF (DAKO), mouse anti-CD63 and rat anti-GATA1 N6 (Santa Cruz Biotechnology), rabbit anti-GATA1 (NF; the antibody was produced against the recombinant N-terminal zinc finger[58]), rabbit anti-GAPDH (14C10, Cell Signaling) and anti-β3 integrin (sc-14009, Santa Cruz Biotechnology). The statistical analysis of the GATA1 data are described in the Supplementary Information.

MPL expression on platelets

The level of MPL protein on the platelet membrane was measured by flow cytometry (Beckman Coulter FC500) using the monoclonal antibodies: APC-labelled IgG1 against CD42b (clone HIP1, BD Pharmingen, 551061), PE-labelled IgG1 against CD110 (clone REA250, Miltenyi Biotec) and a PE-labelled isotype control (clone MOPC-21, BD Pharmingen, 555749). In brief, a sample of EDTA-anticoagulated blood was incubated with anti-CD110 (or control) and anti-CD42b antibodies for 30 min. Mean fluorescence intensity (MFI) produced by the anti-CD110 antibody was measured by flow cytometry on cells gated on the CD42b APC signal, side and forward scatter.

Nanopore sequencing

Oxford Nanopore-based sequencing of long-range PCR-amplified target DNA was performed as previously described[59] with the aim of resolving the genetic architecture of intron 9 of *ITGB3* in a case with Glanzmann's thrombasthenia. The flow cell ran for 3 h, and the mean coverage was 863,986×.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Genotype and phenotype data from the 4,835 participants enrolled in the National Institute for Health Research (NIHR) BioResource for the 100,000 Genomes Project Rare Diseases Pilot can be accessed by application to Genomics England Ltd following the procedure outlined at: https://www.genomicsengland.co.uk/about-gecip/joining-researchcommunity/. The genotype data for the 764 UK Biobank samples will be made available through a data-release process that is being overseen by the UK Biobank (https://www.ukbiobank.ac.uk/). The full blood count data from UK Biobank participants are available from UK Biobank using their access procedures.

The WGS and detailed phenotype data of the remaining 7,348 NIHR BioResource participants can be accessed by application to the NIHR BioResource Data Access Committee (dac@bioresource.nihr.ac.uk). Subject to ethical consent, the genotype data of 6,939 NIHR BioResource participants are also available from the European Genome-phenome Archive (EGA) at the EMBL European Bioinformatics Institute under access procedures managed by EGA. The domain-specific accessions are as follows (refer to the legend of Fig. 1 for domain acronym definitions): BPD, EGAD00001004519; CSVD, EGAD00001004513; EDS, EGAD00001005123; HCM, EGAD00001004514; ICP, EGAD00001004515; IRD, EGAD00001004520; LHON, EGAD00001005122; MPMT, EGAD00001004521; NDD, EGAD00001004522; NPD, EGAD00001004516; PAH, EGAD00001004525; PID, EGAD00001004523; PMG, EGAD00001004517; SMD, EGAD00001004524; SRNS, EGAD00001004518. The ATAC-seq and H3K27ac ChIP–seq data to support the generation of the regulomes are available from GEO (https://www.ncbi.nlm.nih.gov/geo/), EGA (https://ega-archive.org), or referenced to their publication as follows. H3K27ac ChIP–seq: activated CD4^{+} T cells[60], B cells (ERR1043004, ERR1043129, ERR928206, ERR769436), erythroblasts (EGAD00001002377), megakaryocytes (EGAD00001002362), monocytes (ERR829362 (ERS257420), ERR829412 (ERS222466), ERR493634 (ERS214696)), resting CD4^{+} T cells[60]. ATAC-seq: activated CD4^{+} T cells (GSE124867), B cells (SRR2126769 (GSE71338)), erythroblasts (SRR5489430 (GSM2594182)), megakaryocytes (EGAD00001001871), monocytes (EGAD00001006065), resting CD4^{+} T cells (GSE124867). Reported alleles and their clinical interpretation have been deposited with ClinVar under the study names ‘NIHR_Bioresource_Rare_Diseases_13k’, ‘NIHR_Bioresource_Rare_Diseases_Retinal_Dystrophy’, ‘NIHR_Bioresource_Rare_Diseases_MYH9’ and ‘NIHR_Bioresource_Rare_Diseases_PID’. MDT-reported alleles and their clinical interpretation have been deposited in ClinVar (under the name ‘NIHR Bioresource Rare Diseases’) and DECIPHER.

Code availability

Code to run HBASE is available from https://github.com/mh11/VILMAA. The RedPop software package is available from https://gitlab.haem.cam.ac.uk/et341/redpop/.

53. Robinson, P. N. et al. The Human Phenotype Ontology: a tool for annotating and analyzing human hereditary disease. *Am. J. Hum. Genet*. **83**, 610–615 (2008).
54. MacArthur, J. A. et al. Locus Reference Genomic: reference sequences for the reporting of clinically relevant sequence variants. *Nucleic Acids Res*. **42**, D873–D878 (2014).
55. Lek, M. et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).
56. McLaren, W. et al. The Ensembl Variant Effect Predictor. *Genome Biol*. **17**, 122 (2016).
57. Di Michele, M. et al. An integrated proteomics and genomics analysis to unravel a heterogeneous platelet secretion defect. *J. Proteomics* **74**, 902–913 (2011).
58. de Waele, L. et al. Severe gastrointestinal bleeding and thrombocytopenia in a child with an anti-GATA1 autoantibody. *Pediatr. Res*. **67**, 314–319 (2010).
59. Sanchis-Juan, A. et al. Complex structural variants in Mendelian disorders: identification and breakpoint resolution using short- and long-read genome sequencing. *Genome Med*. **10**, 95 (2018).
60. Burren, O. S. et al. Chromosome contacts in activated T cells identify autoimmune disease candidate genes. *Genome Biol*. **18**, 165 (2017).
61. Wijgaerts, A. et al. The transcription factor GATA1 regulates NBEAL2 expression through a long-distance enhancer. *Haematologica* **102**, 695–706 (2017).

Acknowledgements This research was made possible through access to the data and findings generated by two pilot studies for the 100,000 Genomes Project. The enrolment was coordinated for one by the NIHR BioResource and for the other by Genomics England Ltd (GEL), a company wholly owned by the Department of Health in the United Kingdom. These pilot studies were mainly funded by grants from the NIHR in England to Cambridge University Hospitals and GEL, respectively. Additional funding was provided by the British Heart Foundation (BHF), MRC, NHS England, the Wellcome Trust and many other fund providers (see funding acknowledgements for individual researchers). The pilot studies use data provided by patients and their close relatives and collected by the NHS and other healthcare providers as part of their care and support. The vast majority of participants in the two pilot studies have been enrolled in the NIHR BioResource. We thank all volunteers for their participation and the NIHR Biomedical Research Centres (BRC), NIHR BioResource Centres, NHS Trust Hospitals, NHS Blood and Transplant and their staff for their contribution. This research has been conducted using the UK Biobank resource under Application Number 9616, granting access to DNA samples and accompanying participant data. UK Biobank has received funding from the MRC, Wellcome Trust, Department of Health, BHF, Diabetes UK, Northwest Regional Development Agency, Scottish Government and Welsh Assembly Government. The MRC and Wellcome Trust had a key role in the decision to establish the UK Biobank. A. McMahon and J. Morales are funded by The Wellcome Trust (WT200990/Z/16/Z) and the European Molecular Biology Laboratory; K.G.C.S. holds a Wellcome Investigator Award, MRC Programme Grant (number MR/L019027/1); M.I.M. is a Wellcome Senior Investigator and receives support from the Wellcome Trust (090532, 0938381) and is a member of the DOLORisk consortium funded by the European Commission Horizon 2020 (ID633491); R. Horvath is a Wellcome Trust Investigator (109915/Z/15/Z), who receives support from the Wellcome Centre for Mitochondrial Research (203105/Z/16/Z), MRC (MR/N025431/1), the European Research Council (309548), the Wellcome Trust Pathfinder Scheme (201064/Z/16/Z), the Newton Fund (UK/Turkey, MR/N027302/1) and the European Union H2020 – Research and Innovation Actions (SC1-PM-03-2017, Solve-RD); D.L.B. is a Wellcome clinical scientist (202747/Z/16/Z) and is a member of the DOLORisk consortium funded by the European Commission Horizon 2020 (ID633491); J.S.W. is funded by Wellcome Trust [107469/Z/15/Z], NIHR Cardiovascular Biomedical Research Unit at Royal Brompton & Harefield NHS Foundation Trust and Imperial College London; A.J.T. is supported by the Wellcome Trust (104807/Z/14/Z) and the NIHR Biomedical Research Centre at Great Ormond Street Hospital for Children NHS Foundation Trust and University College London; L. Southgate is supported by the Wellcome Trust Institutional Strategic Support Fund (204809/Z/16/Z) awarded to St George’s, University of London; M.J.D. receives funding from Wellcome Trust (WT098519MA); M.C.S. holds an MRC Clinical Research Training Fellowship (MR/R002363/1); J.A.S. is funded by MRC UK grant MR/M012212/1; A.J.M. received funding from an MRC Senior Clinical Fellowship (MR/L006340/1); C. Lentaigne received funding from an MRC Clinical Research Training Fellowship (MR/J011711/1); M.R.W. holds a NIHR award to the NIHR Imperial Clinical Research Facility at Imperial College Healthcare NHS Trust; C. Williamson holds a NIHR Senior Investigator Award; M. A. Kurian holds a NIHR Research Professorship (NIHR-RP-2016-07-019) and Wellcome Intermediate Fellowship (098524/Z/12/A); M.J.C. is an NIHR Senior Investigator and is funded by the NIHR Barts Biomedical Research Centre; N. Cooper is partially funded by NIHR Imperial College Biomedical Research Centre; C. Hadinnapola was funded through a PhD Fellowship by the NIHR Translational Research Collaboration - Rare Diseases; A.D.M. and S.K.W. were funded by the NIHR Bristol Biomedical Research Centre; E.L.M. received funding from the NIHR Biomedical Research Centre at University College London Hospitals; K.C.G. received funding from the NIHR Great Ormond Street Biomedical Research Centre; I.R. and E. Louka are supported by the NIHR Translational Research Collaboration - Rare Diseases; J.C.T., J.M.T. and S. Patel are funded by the NIHR Oxford Biomedical Research Centre; G. Arno is funded by a Fight for Sight (UK) Early Career Investigator Award (5045-5046); All authors affiliated with Moorfields Eye hospital and Institute of Ophthalmology are funded by the NIHR Moorfields Biomedical Research Centre and UCL Institute of Ophthalmology, Fight for Sight (UK) Early Career Investigator Award, Moorfields Eye Hospital Special Trustees, Moorfields Eye Charity, Foundation Fighting Blindness (USA) and Retinitis Pigmentosa Fighting Blindness; A.T.M. is funded by Retinitis Pigmentosa Fighting Blindness, P.Y.-W.-M. is supported by grants from MRC UK (G1002570), Fight for Sight (1570/1571 and 24TP171), NIHR (IS-BRC-1215-20002); S.O.B. is supported by NIHR Translational Research Collaboration - Rare Diseases (01/04/15-30/04/2017); A.R.W. works for the NIHR Moorfields Biomedical Research Centre and the UCL Institute of Ophthalmology and Moorfields Eye Hospital; the following NIHR Biomedical Research Centres contributed to the enrolment for the ICP domain: Imperial College Healthcare NHS Trust, Guy’s and St Thomas’ NHS Foundation Trust and King’s College London. All authors affiliated with Moorfields Eye hospital and Institute of Ophthalmology are funded by the NIHR Biomedical Resource Centre at UCL Institute of Ophthalmology and Moorfields; A.C.T. is a member of the International Diabetic Neuropathy Consortium, the Novo Nordisk Foundation (NNF14SA0006) and is a member of the DOLORisk consortium funded by the European Commission Horizon 2020 (ID633491); J. Whitworth is a recipient of a Cancer Research UK Cambridge Cancer Centre Clinical Research Training Fellowship; S.A.J. is funded by Kids Kidney Research; D.P.G. is funded by the MRC, Kidney Research UK and St Peters Trust for Kidney, Bladder and Prostate Research; The MPGN/DDD/C3 Glomerulopathy Rare Disease Group contributed to the recruitment and analysis for the PMG domain; K.J.M. is supported by the Northern Counties Kidney Research Fund; P.H.D. receives funding from ICP Support; T.K.B. received a PhD fellowship from the NHSBT and British Society of Haematology; H.S.M. receives support from BHF Programme Grant RG/16/4/32218; A.L. is a BHF Senior Basic Science Research Fellow - FS/13/48/30453; K.F. and C.V.-G. are supported by the Research Council of the University of Leuven (BOF KU Leuven, Belgium; OT/14/098); H.J.B. works for the Netherlands CardioVascular Research Initiative (CVON); Fiona Cunningham, Aoife McMahon, Glen Threadgold, and Joannella Morales received funding from the Wellcome Trust (grant numbers WT108749/Z/15/Z and WT200990/Z/16/Z) and the European Molecular Biology Laboratory. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, the Department of Health and Social Care of England or any of the other funding agencies.

Author contributions Details of author contributions can be found in the Supplementary information, which contains the full list of consortium members and working groups.

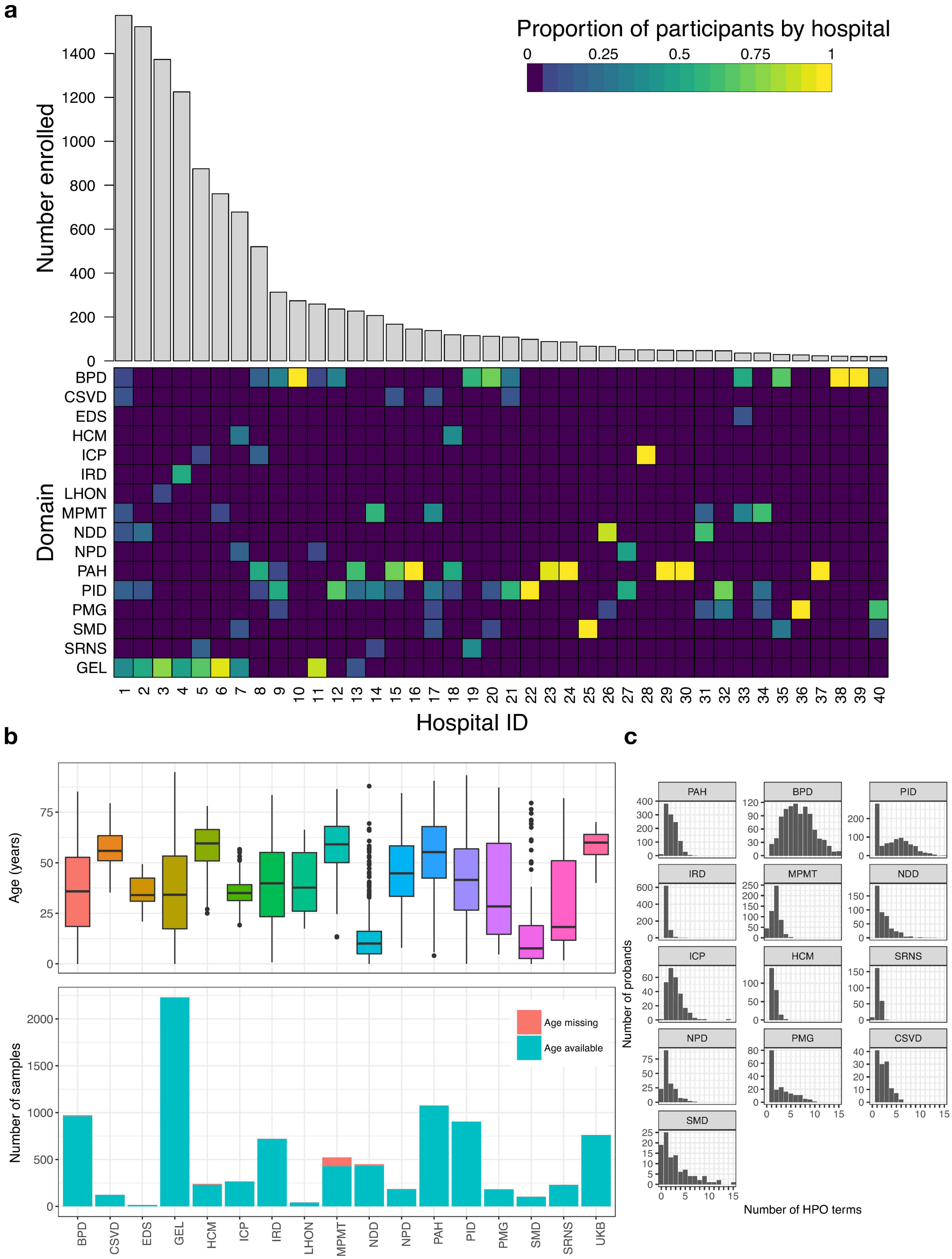
Competing interests A.M.K. had no competing interests at the time of the study, but after the study received an educational grant from CSL Behring to attend the ISTH meeting (2017); T.J.A. has received consultancy payments from AstraZeneca within the past 5 years and has received speaker honoraria from Illumina; A. Rogers, C. Cheah, C. Steward, E.B., K. Tate, N. Lench and R. Prathalingam are employees of Congenica; B. Tolhuis, J. Findhammer, J.K., M.V. and T. Karten are employees of GENALICE; C. Colombo, C. Geoghegan, C.J.B., C. Rees, D.R.B., J.F.P., J. Hughes, R.J.G., S. Humphray, S. Hunter and T.S.A.G. are employees of Illumina Cambridge; C.V.-G. is the holder of the Bayer and Norbert Heimburger (CSL Behring) Chair; K.J.M. previously received funding for research and is currently on the scientific advisory board of Gemini Therapeutics; M.C.S. received travel and accommodation fees from NovoNordisk; D.M.L. serves on advisory boards for Agios, Novartis and Cerus; M.I.M. serves on advisory panels for Pfizer, NovoNordisk and Zoe Global, has received honoraria from Pfizer, NovoNordisk and Eli Lilly, has stock options in Zoe Global, and has received research funding from Abbvie, AstraZeneca, Boehringer Ingelheim, Eli Lilly, Janssen, Merck, NovoNordisk, Pfizer, Roche, Sanofi Aventis, Servier and Takeda; DLB has acted as a consultant on behalf of Oxford Innovation in the last 2 years for the following companies: Amgen, CODA therapeutic, Bristows, Lilly, Mundipharma, Regeneron and Theranexus, he holds an MRC Industrial Partnership grant with Astra Zeneca. The remaining authors declare no competing interests.

Additional information

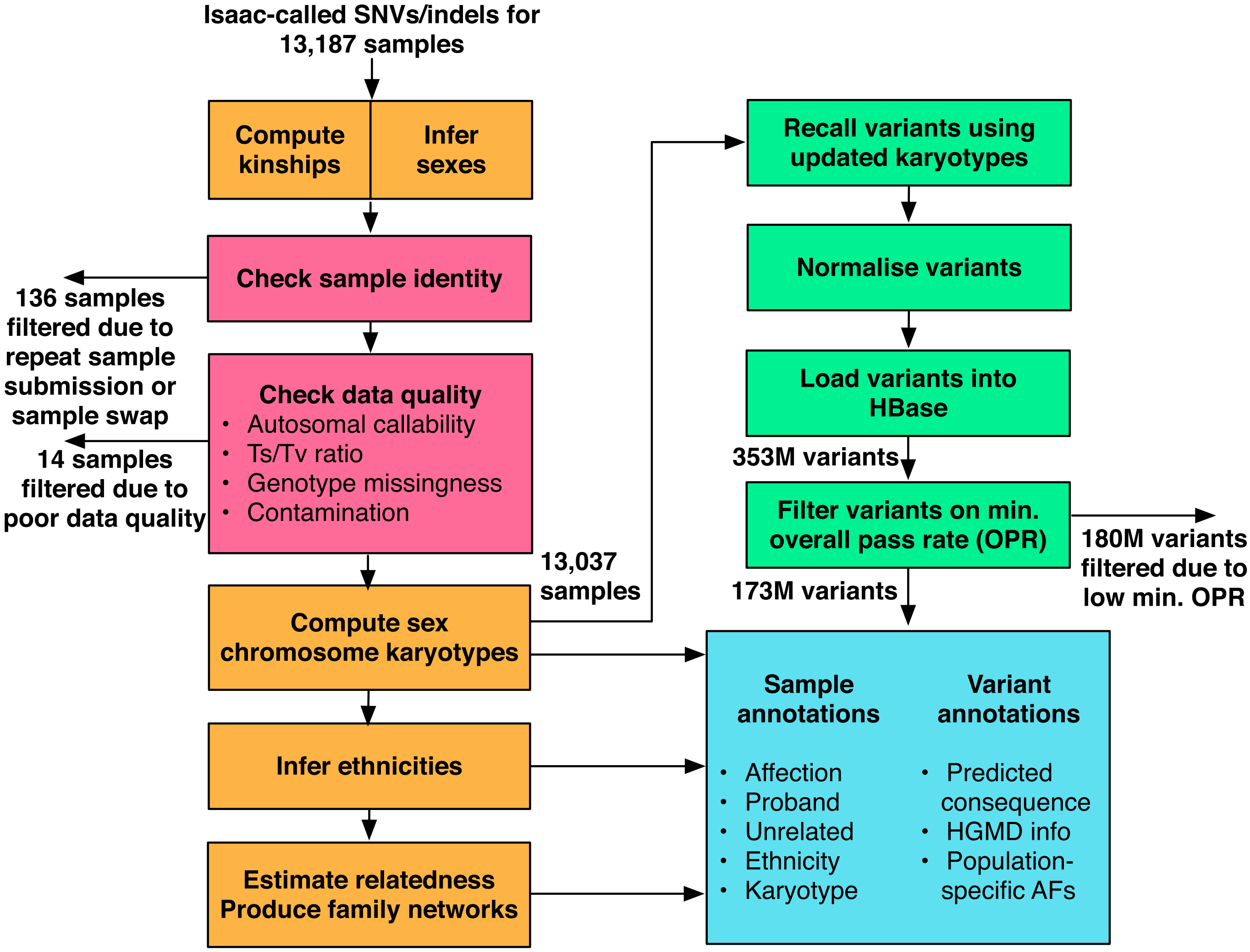
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2434-2.

Correspondence and requests for materials should be addressed to E.T., F.L.R. or W.H.O.

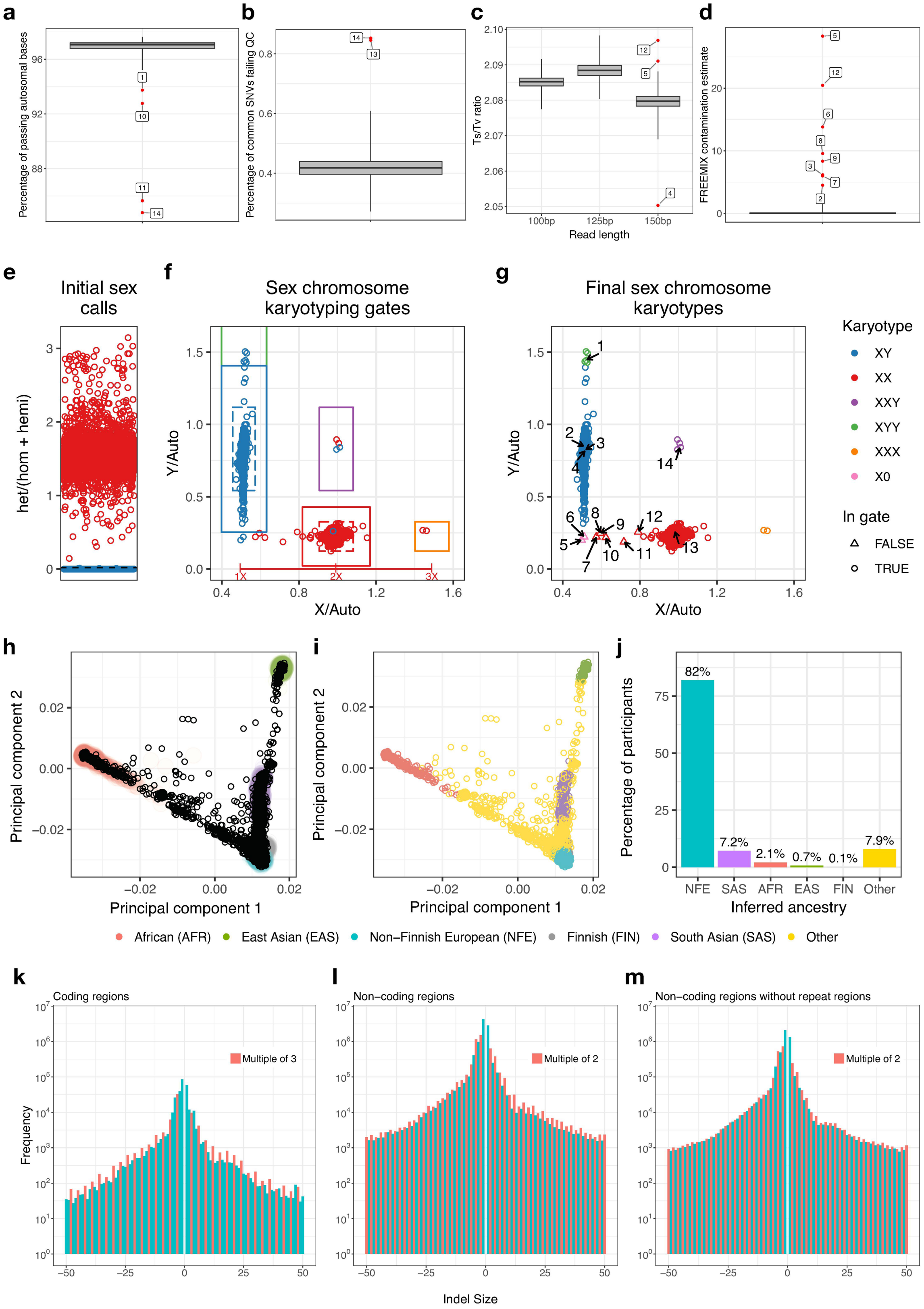
Peer review information *Nature* thanks Heidi L. Rehm, V. G. Sankaran, Shamil Sunyaev and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



Extended Data Fig. 1 | Demographic and phenotypic characteristics. **a**, The number of enrolments at the 40 hospitals with at least 20 enrolled participants. The heat map shows the distribution of enrolments over domains at each of the 40 hospitals. Hospital IDs are described in Supplementary Table 1. **b**, Top, age at recruitment for all probands in the 15 rare disease domains, GEL and UKB. Bottom, counts of probands in each domain with and without an available age at recruitment. **c**, Histograms of the number of HPO terms appended to affected probands for 13 of the rare disease domains.

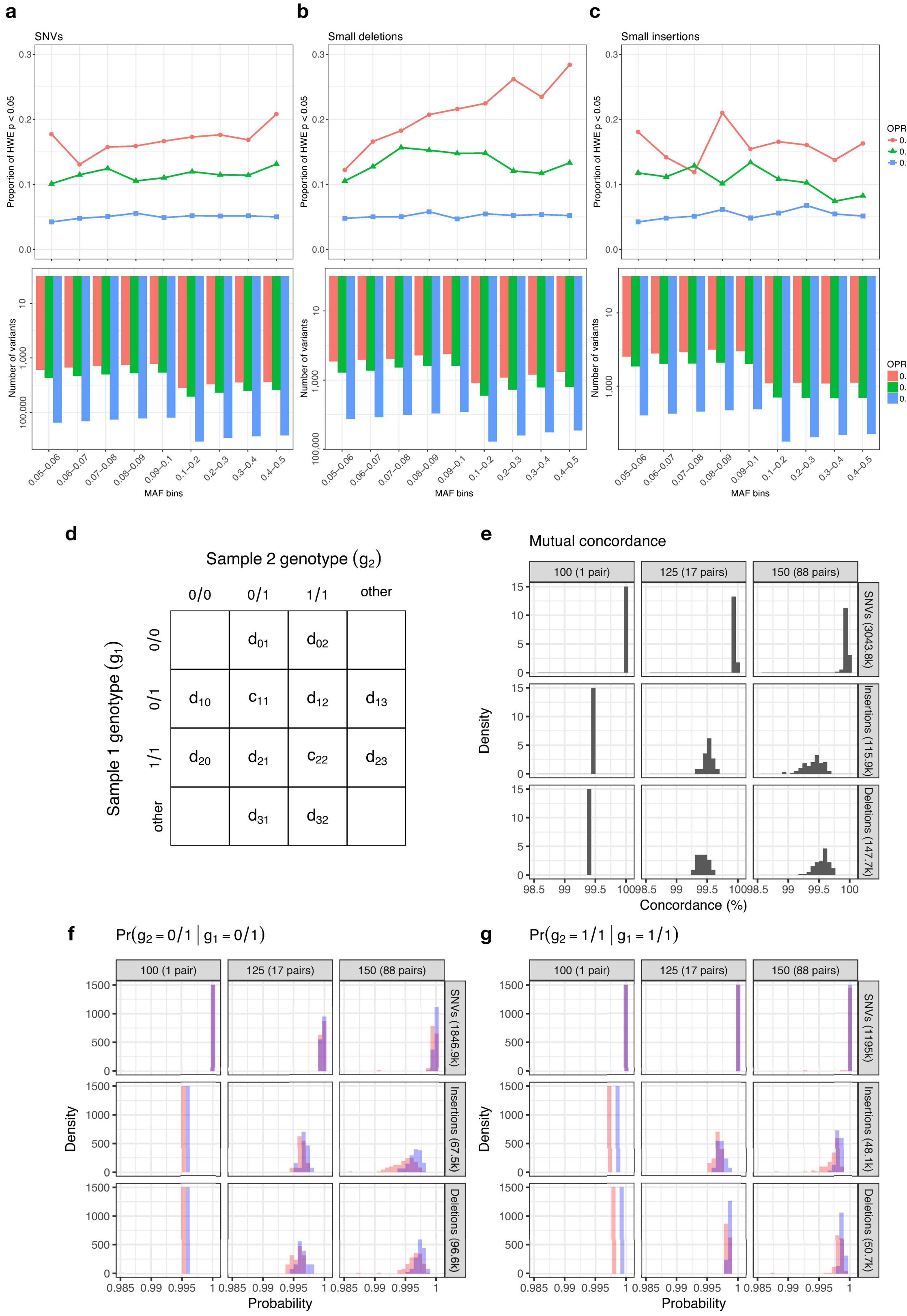


Extended Data Fig. 2 | Flowchart of the bioinformatic data processing. Flowchart describing the processing of samples and variants. Beginning at the top left, all samples were checked for data quality (Extended Data Fig. 3). Quick kinship and sex checks were regularly performed to ensure consistency with reported sex and family information. Samples that failed quality control, samples with clearly discordant sex data and the sub-optimal replicates of repeated samples were removed before further analysis (pink boxes). Sex chromosome karyotypes, ethnicities and relatedness/family trees were computed on these filtered samples (orange boxes) and variants were recalled for those samples with X/Y-chromosome ploidies different to those automatically predicted by the quick checks. After variant normalization, variant calls were loaded into HBase and merged, and summary statistics were calculated, stratified by technical factors (100, 125 and 150 bp) and ancestry (for example, African) (green boxes). Variant-specific minimum overall pass rates were calculated and used to filter inaccurately genotyped variants (Extended Data Fig. 4). Finally, variants were annotated in HBase with predicted consequence information and information from external databases, including allele frequencies (AF) (for example, gnomAD) (blue box).



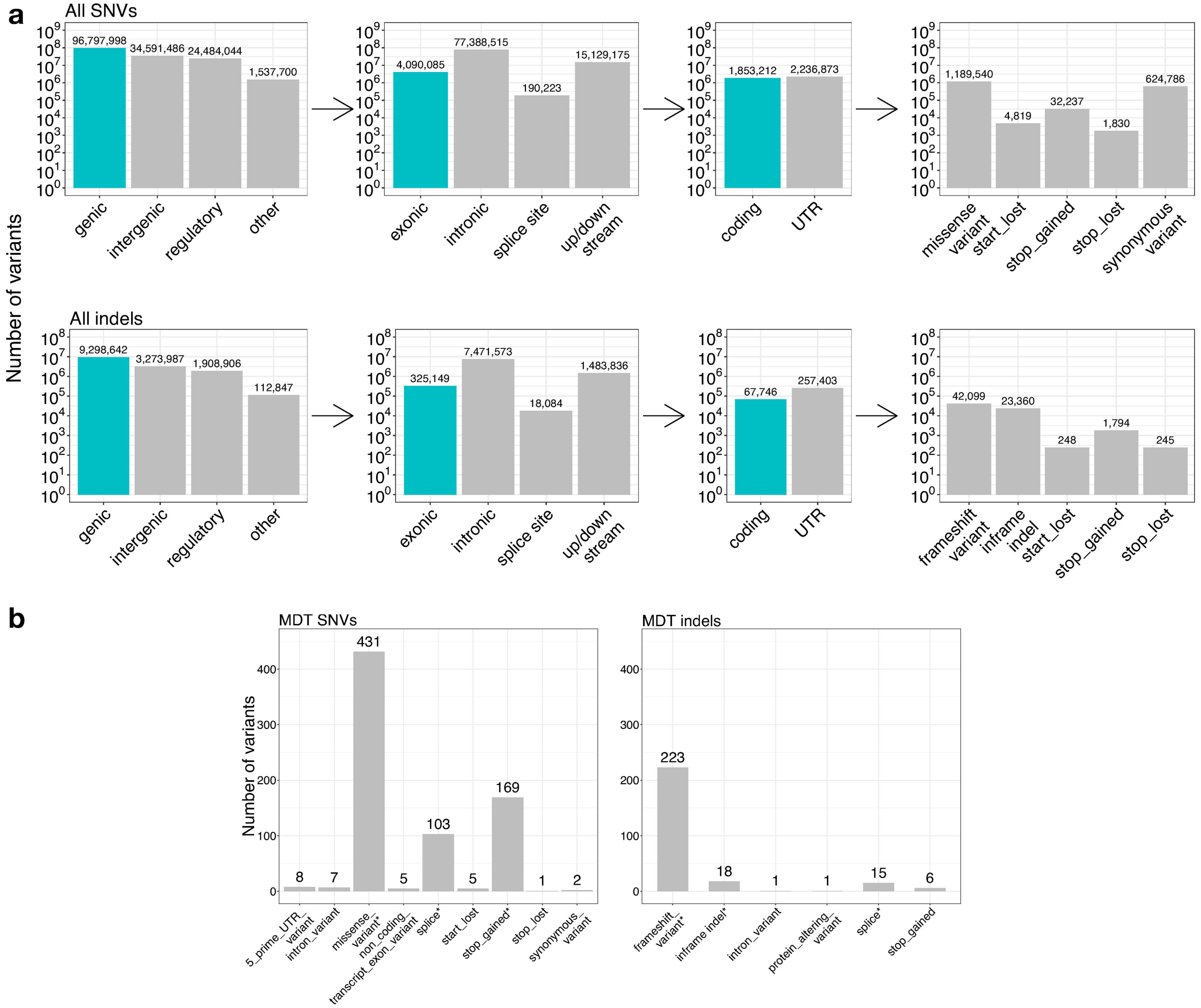
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Sample quality control, sex chromosome karyotyping and ancestry inference. a, The percentage of quality-control-passing autosomal bases ($n = 13{,}187$; 4 exclusions highlighted). **b**, The percentage of common SNVs that failed quality control ($n = 13{,}187$; 2 exclusions highlighted). **c**, Batch-specific box plots of Ts/Tv ratios ($n = 377$ for 100-bp samples; $n = 3{,}154$ for 125-bp samples; $n = 9{,}656$ for 150-bp samples; 3 exclusions highlighted). **d**, FREEMIX values representing sample contamination ($n = 13{,}187$; 8 exclusions highlighted). **a**–**d**, Excluded samples are marked in red and labelled with an integer. Three samples were excluded because they failed more than one of the four quality control checks (samples 5, 12 and 14). The centre line of each box plot indicates the median and the lower and upper hinges indicate the 25th and 75th percentiles, respectively. The vertical line of each boxplot extends to 1.5× the interquartile range from each hinge. **e**, The number of heterozygous variants divided by the number of homozygous and hemizygous variants coloured by the initial predicted sexes for 13,037 samples. **f**, Scatter plot of ratios of X/Auto and Y/Auto coloured by the initial sex calls and showing the five sex karyotyping gates. **g**, Scatter plot of ratios of X/Auto and Y/Auto coloured by the final sex chromosome karyotype. Circles indicate samples falling within a sex karyotyping gate and triangles indicate samples falling outside all sex karyotyping gates. 1, confirmed XYY case; 2–4, confirmed XY female cases; 5, 6, confirmed XO cases; 7, confirmed XO case, this sample has some part of the second X chromosome present; 8–10, samples with a large part of the X chromosome missing; 11–12, samples with multiple deletions on the X chromosome; 13, sample with two almost identical X chromosomes (normal karyotype); 14, confirmed XXY case. **h**, Projection of the 13,037 samples, shown as circles, onto the 1000 Genomes-derived PCAs. The 1000 Genomes samples are shown as diffuse points underneath in colour. **i**, Projection of the 13,037 samples, shown as circles, coloured by assigned population. **j**, The number of individuals assigned to each population. The percentages are shown above each bar. NFE, Non-Finnish European; SAS, South Asian; AFR, African; EAS, East Asian; FIN: Finnish. **k**–**m**, Distribution of the sizes of small insertions (indel size > 0) and small deletions (indel size < 0) in coding regions (**k**), non-coding regions (**l**) and non-coding regions excluding repetitive regions, specifically, the RepeatMasker track from the UCSC table browser and the Tandem Repeats Finder locations from the UCSC hg19 full dataset download (**m**). In coding regions, natural selection against frameshift variants results in a systematic depletion of indel sizes that are not a multiple of 3 bp. In non-coding regions, there is a slight excess of indel sizes that are a multiple of 2 bp, but this pattern is almost indiscernible if repetitive regions are excluded.

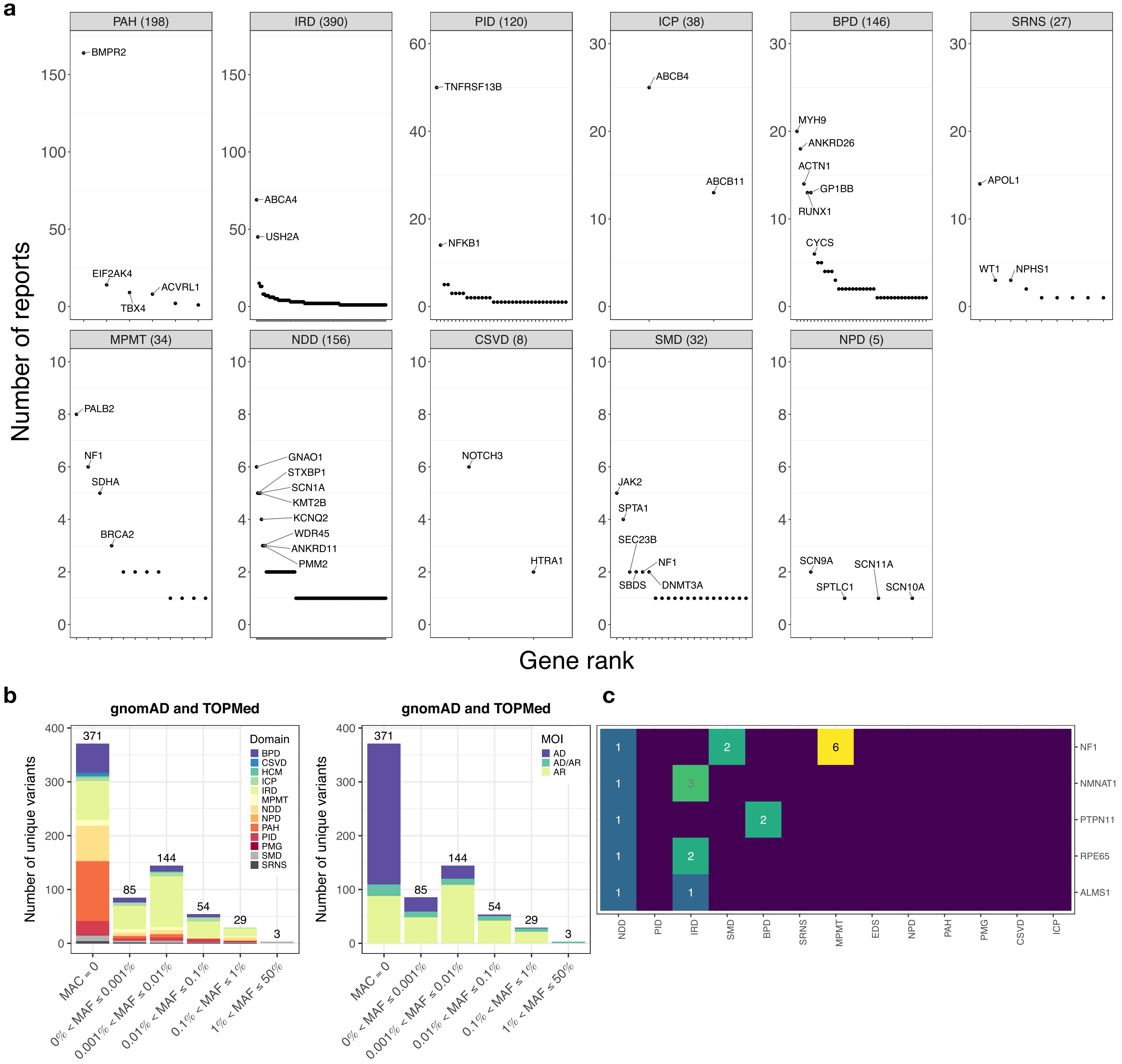


Extended Data Fig. 4 | See next page for caption.

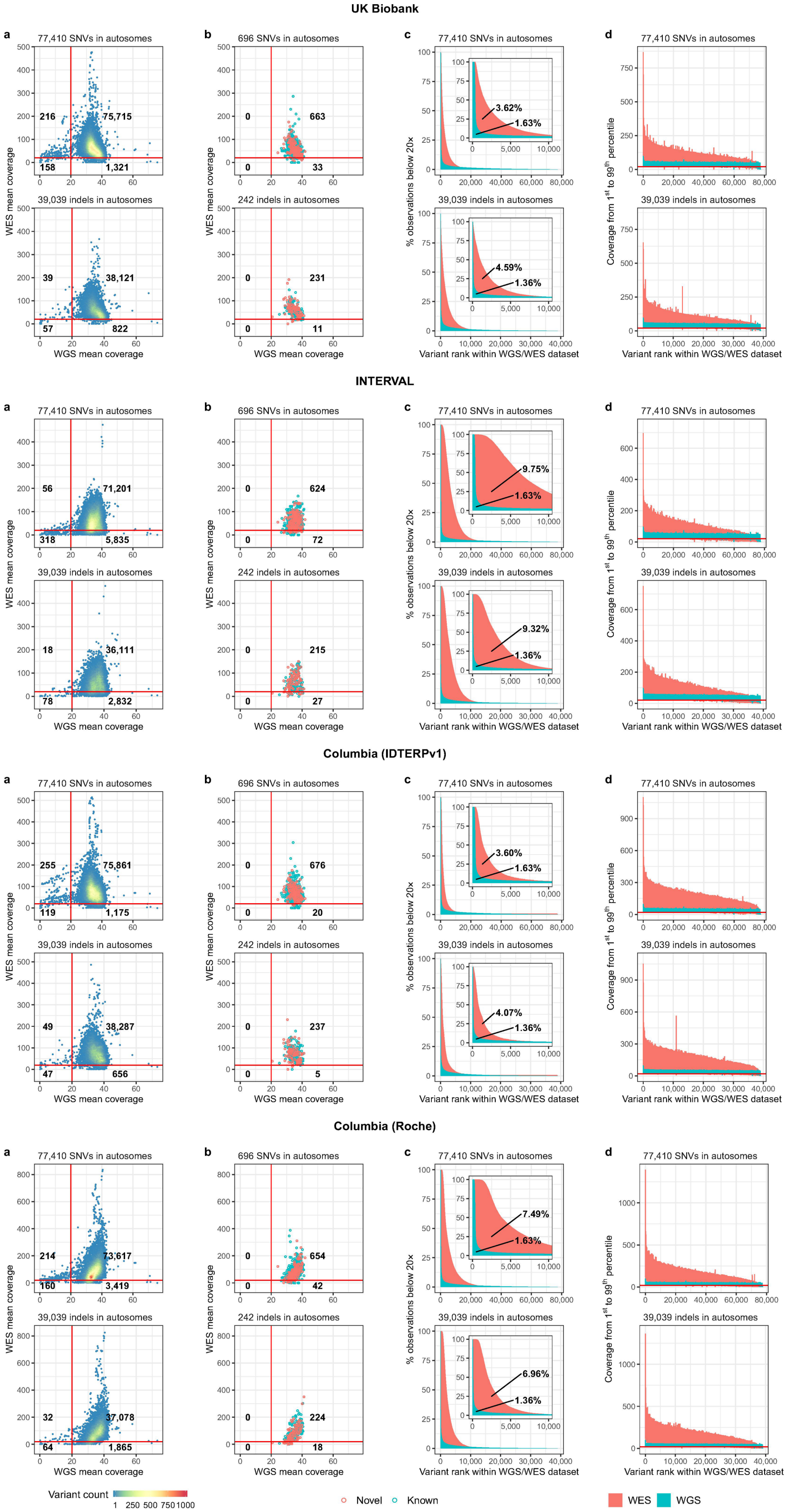
Extended Data Fig. 4 | Variant quality control. a–**c**, The proportion of *P* values computed to test the null hypothesis of Hardy–Weinberg equilibrium < 0.05 among 8,510 unrelated Europeans across different allele frequency (AF) bins for SNVs (**a**), small deletions (**b**) and small insertions (**c**). The number of variants in each overall pass rate (OPR) and allele frequency bin are shown in the bottom sub-panels. **d**, Table showing the possible combinations of genotypes in a pair of samples. The variables in the cells represent numbers of variants (see Supplementary Information for use). **e**–**g**, Three measures of genotype concordance (Supplementary Information) for pairs of duplicates and twins with results from 100-, 125- and 150-bp reads shown from left to right. **e**, Distribution of mutual non-reference concordance in pairs of duplicates and twins. **f**, Probability of having a heterozygous genotype in a sample, given its duplicate or twin has this heterozygous genotype. **g**, Probability of having a non-reference homozygous genotype in a sample, given its duplicate or twin has this homozygous genotype. In **e**–**g**, the mean number of variants of each type used to compute concordance is shown in brackets after the variant type label. In **f**, **g**, red and blue colours represent the distribution of the lowest and highest of the two probabilities (sample 1 compared to sample 2 and sample 2 compared to sample 1) in a pair of duplicates or twins.



Extended Data Fig. 5 | Breakdown of genetic variants by their predicted primary consequence. a, Counts of SNVs and indels in various Variant Effect Predictor consequence classes shown on logarithmic scales with exact numbers above each bar. Variants in the turquoise bars are subdivided into more granular regions of genome space in the following panel in a recursive manner from left to right. Categories have been chosen to represent the most severe transcriptional consequences at each stage: that is, from left, overall genome space, within genes, exonic parts of genes and protein-coding regions. **b**, Count of MDT SNVs and indels in various consequence classes with exact numbers above each bar. An asterisk denotes a supercategory with ‘missense_variant’ including ‘missense_variant’ or ‘missense_variant & splice_region_variant’; ‘splice’ including ‘splice_acceptor_variant’, ‘splice_donor_variant’, ‘splice_donor_variant & coding_sequence_variant’ or ‘splice_region_variant’ or ‘splice_region_variant & intron_variant’; ‘stop_gained’ including ‘stop_gained’, ‘stop_gained & splice_region_variant’ or ‘stop_gained & splice’; ‘frameshift variant’ including ‘frameshift_variant’, ‘frameshift_variant & splice_region_variant’ or ‘retained_intron’; ‘inframe indel’ including ‘inframe_deletion’ or ‘inframe_insertion’.

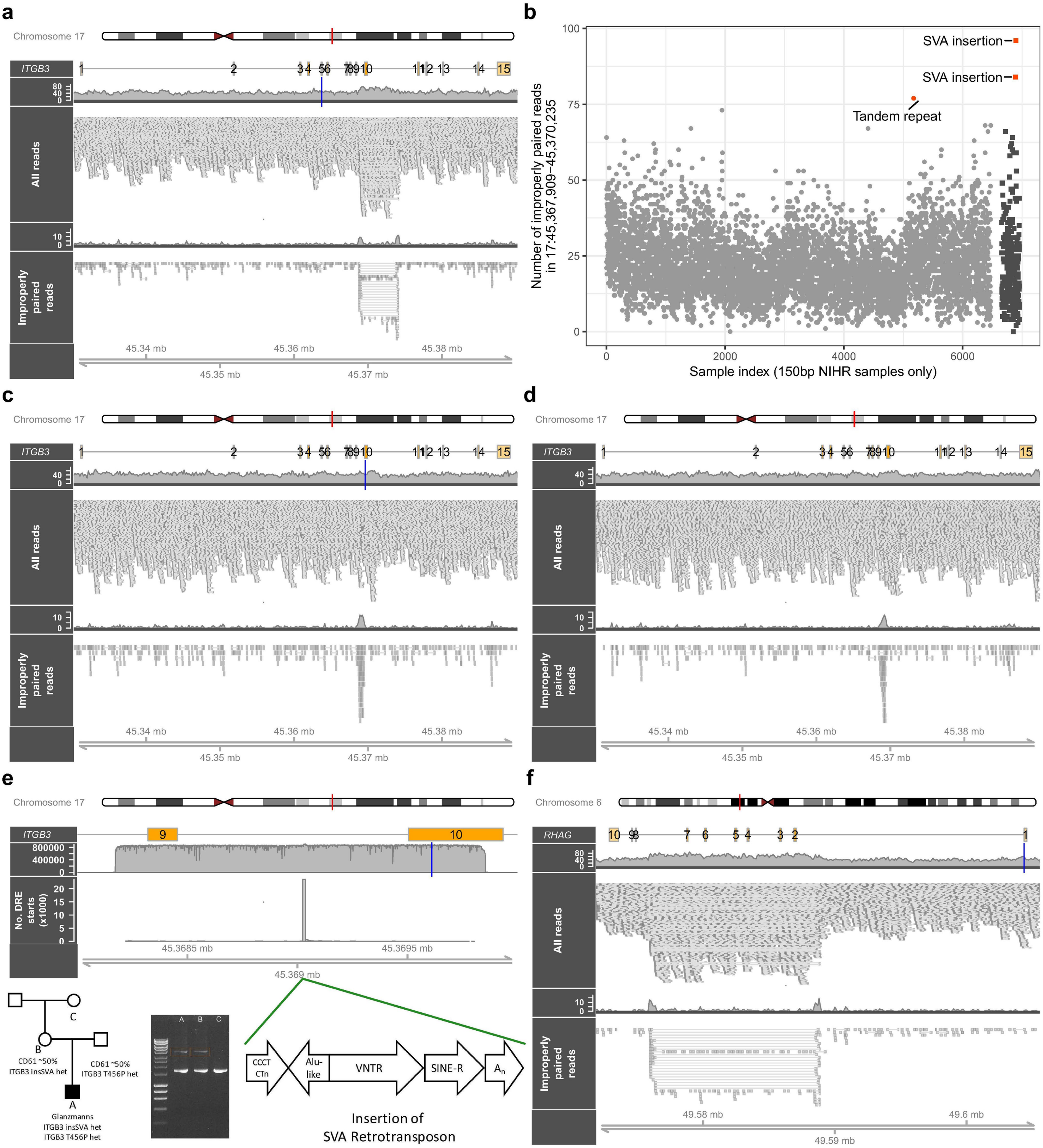


Extended Data Fig. 6 | Breakdown of diagnostic reports by domain. **a**, Number of reports issued for the 11 rare disease domains that issued clinical reports. Each panel corresponds to a domain, the title denotes the domain acronym and number of reports issued. PMG and EDS domains are not shown because no reports were issued for cases in these domains. The panels are arranged in decreasing order of the maximum number of within domain reports issued for a single DGG. Each point represents a gene featuring in at least one report for a case in the domain. The genes with the most reports issued for each domain are labelled. Full details of all the reports issued are given in Supplementary Table 2. **b**, The number of distinct reported autosomal short variants (SNVs and indels) for each domain in different gnomAD/TOPMed allele frequency bins in samples of European ancestry, broken down by rare disease domain (left) and by mode of inheritance (right). The domain acronyms are defined in Supplementary Table 1. MOI, mode of inheritance; AD, autosomal dominant; AR, autosomal recessive. For a given position and minor allele, the combined MAF was defined as the sum of allele counts divided by the sum of allele numbers over gnomAD and TOPMed. The first bin in the plots (MAC = 0) corresponds to variants not observed in either gnomAD or TOPMed. **c**, Some genes featured in reports for cases in more than one domain. The heat map shows the number of reports featuring these genes, broken down by domain.

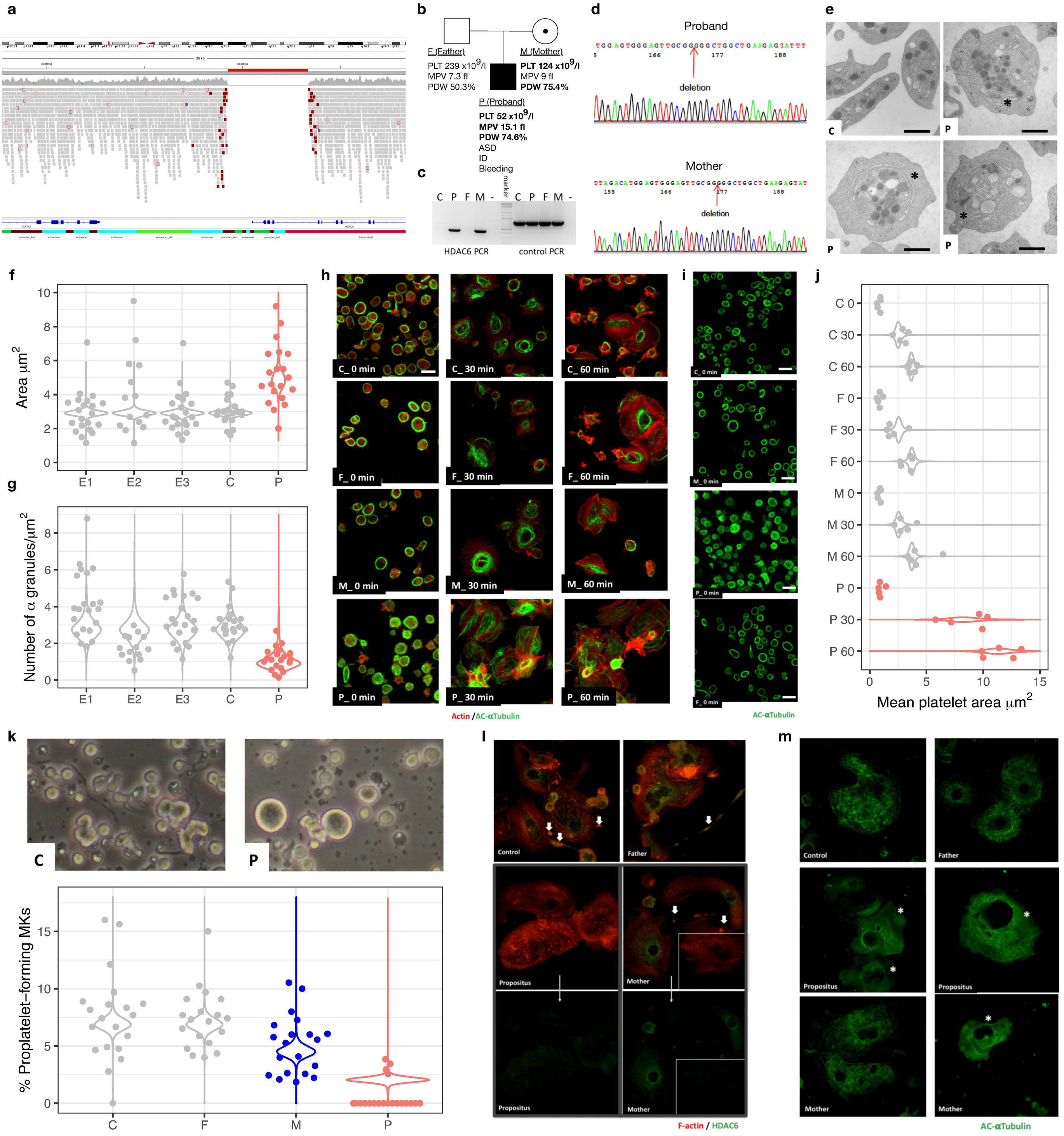


Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Comparison of WGS and WES for genetic testing. **a**–**d**, For each of four WES datasets—'UK Biobank', 'INTERVAL', 'Columbia (IDTERPv1)' and 'Columbia (Roche)'—four groups of panels are shown, each of which corresponds to a different comparison of coverage characteristics, as follows. **a**, WGS versus WES mean coverage at 116,449 sites of diagnostic importance (Supplementary Information). The red axes show the threshold for clinical reporting and the numbers of variants in each quadrant are indicated. **b**, WGS versus WES coverage of the MDT-reported known (turquoise) and novel (salmon) SNVs and indels in autosomal diagnostic-grade genes. **c**, The percentage of samples with coverage below the threshold for clinical reporting, with variants ranked on the *x*-axis by their corresponding values on the *y*-axis within the WGS and WES datasets. The bar plots corresponding to WGS are superimposed on those corresponding to WES. The inset shows the mean percentage of individuals covered below 20× by WGS and WES in a magnified view. **d**, Vertical bars indicate the 1–99% coverage range in WGS (turquoise) and WES (salmon), with variants ranked by the mean coverage values within the WGS and WES data sets.

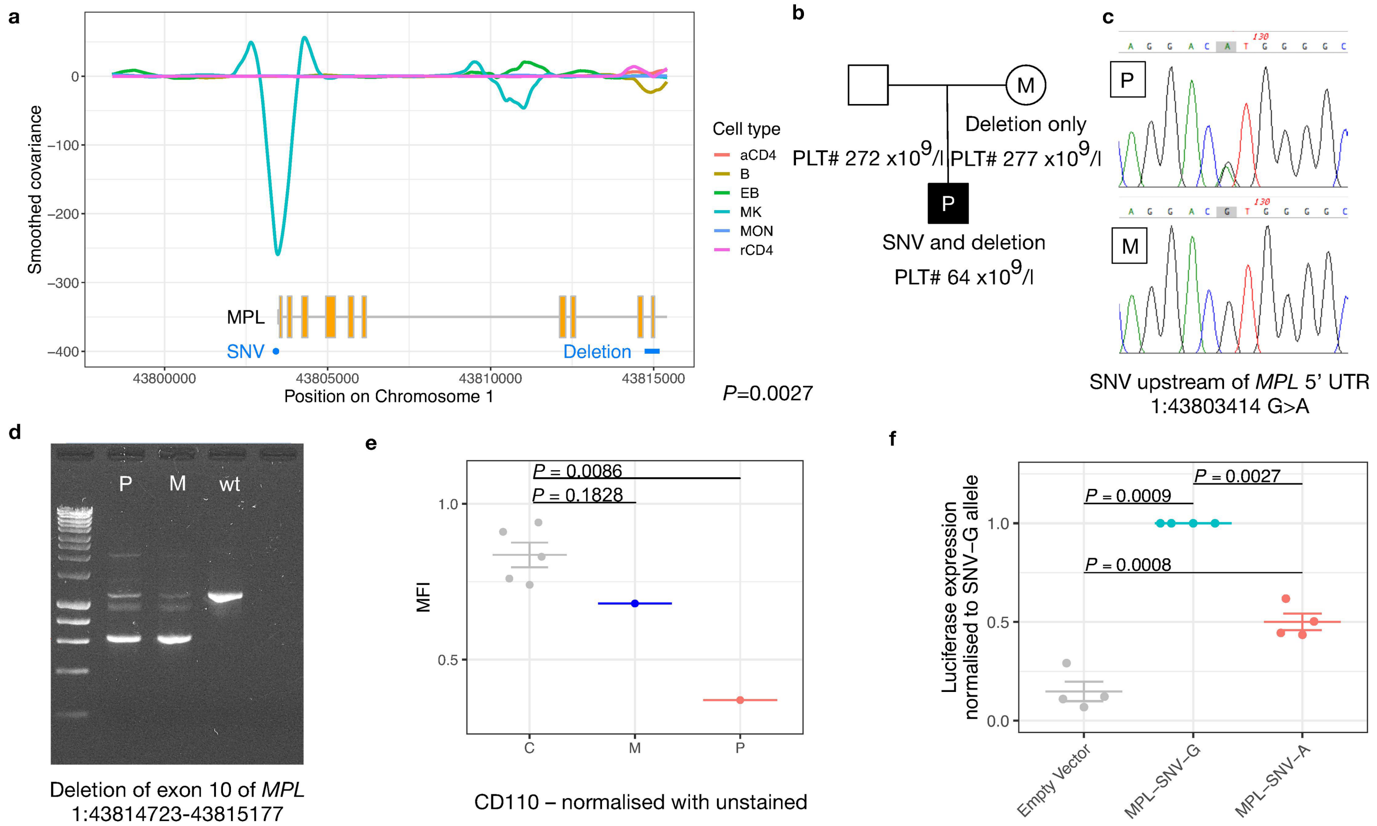


Extended Data Fig. 8 | Cases with protein-null phenotypes. a, Alignments in the *ITGB3* locus for an individual with Glanzmann's thrombasthenia with a premature stop (blue bar) and a tandem repeat revealed by improperly mapped read pairs. **b**, Number of improperly mapped read pairs in the ninth intron of *ITGB3* in 6,656 samples sequenced by 150-bp reads before (light grey dots) or after (dark grey squares) the data freeze. The patients with Glanzmann's thrombasthenia with the tandem repeat and with the SVA insertion, and the carrier mother of the latter, are highlighted. **c**, **d**, Alignments in the *ITGB3* locus for the proband with Glanzmann's thrombasthenia (**c**) and his mother (**d**) with a p.T456P variant for the proband (blue bar) and an insertion revealed by an excess of mapped reads for the ninth intron for the proband and his mother. **e**, Top, long-read alignments for the PCR-amplified *ITGB3* DNA from the proband with Glanzmann's thrombasthenia covering the element with excess reads. Downstream read element (DRE) starts are represented in the histogram. Bottom (from left to right), the pedigree for the patient with Glanzmann's thrombasthenia (A, proband; B, mother; C, grandmother) with the flow cytometry measurements of platelet GPIIbIIIa expression indicated as the percentage of normal levels and genotypes; confirmation of the insertion by gel electrophoresis of PCR products covering the insertion; diagram of the inserted SVA retrotransposon element (insSVA). **f**, Alignments in the *RHAG* locus of the Rh-null case with a splice donor variant (blue bar) and a tandem duplication revealed by improperly mapped read pairs.



Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Deletion of a *GATA1* enhancer and part of the *HDAC6* open-reading frame and its effects. **a**, WGS reads show a hemizygous 4,108-bp deletion (X: 48,659,245–48,663,353) in the proband. **b**–**k**, P, proband; F, father; M, mother; C, control. **b**, Pedigree of the proband with thrombocytopenia and autism. PLT, platelet count; MPV, mean platelet volume; PDW, platelet distribution width; ASD, autism spectrum disorder; ID, intellectual disability. **c**, Left, representative image of $n = 2$ rounds of gel electrophoresis showing presence and absence of short PCR amplicons using primers flanking the deletion. Right, control PCR. '-', no DNA added. **d**, Sanger sequencing of PCR fragments (shown in **c**) with primers flanking the 4,801-bp deletion. The red arrow points to the position of the fusion between base pair 48,659,245 and base pair 48,663,353. **e**, Electron microscopy images ($n = 1$ sample preparation per subject) show that platelets of the proband were larger and rounder than those of the control (unrelated healthy control), and in some instances had abnormal semi-circular empty vacuoles (marked by an asterisk) and a depletion of alpha granules. Scale bars, 1.5 µm. **f**, **g**, Analysis of electron microscopy images ($n = 21, 14, 21, 20$ and 20 platelets in samples E1, E2, E3, C and P, respectively); E1, E2, E3 and C are controls; the data for E1, E2 and E3 were obtained from a previous study[61]. Dot plots of platelet area (µm^2) and the alpha granule count per unit area (µm^{-2}), computed using ImageJ. The underlying violin plots show posterior predictive densities for the mean platelet area or granule density in controls and in the proband under a mixed model accounting for intra-individual correlation. The 90% credible intervals for the ratio of the mean in the proband to the mean in controls were 1.38–2.03 and 0.15–0.87 for area and granule density, respectively. The abnormalities of platelet area and alpha granule density in the proband are very similar to the defects described in *GATA1* deficiency[61]. **h**, Platelet spreading analysis using SIM (*Z*-stacks) and staining for F-actin (red) and acetylated α-tubulin (green). Washed platelets were spread on fibrinogen for 0 (basal condition), 30 and 60 min for control, father, mother and proband. This experiment was performed once and representative images are shown. Scale bars, 1.5 µm. **i**, Platelet analysis using SIM and staining for acetylated α-tubulin (green) before spreading (time point 0). The microtubule marginal bands are clearly disturbed and hyper-acetylated for non-activated platelets of the proband; whereas those of the father and mother are normal. This experiment was performed once. Scale bars, 1.5 µm. **j**, Dot plots of the mean ImageJ-quantified platelet area in groups of $n = 5$ images of F-actin-stained platelets at three time points (0, 30 and 60 min after spreading on fibrinogen) for the control, father, mother and proband. There was no evidence of a difference between the mean of the mean platelet area of either the father or the mother and the control within time points ($P > 0.12$ for all six two-sided Welch *t*-tests), so the father and mother were treated as controls in subsequent modelling. The underlying violin plots show posterior predictive densities for the mean platelet area at time points 30 and 60 min under a mixed model accounting for intra-individual correlation. The 90% credible intervals for the ratio of the mean in the proband to the mean in controls were 1.87–4.56 and 2.07–3.61 at time points 30 and 60 min, respectively. **k**, Top, representative images from the control and the proband. In the latter, large megakaryocytes are present but proplatelet formation is strongly reduced. Bottom, the quantification of proplatelet formation by megakaryocytes at day 12 of differentiation from cultures performed in duplicate for each individual. Ten images per culture were used to compute the percentage proplatelet-forming megakaryocytes per individual, shown as dot plots. There was no evidence of a difference in the mean of the percentage between the father and the control ($P = 0.90$, two-sided Welch *t*-test), so the father was treated as a control in subsequent modelling. The underlying violin plots show posterior predictive densities for the percentage proplatelet-forming megakaryocytes in controls, in the mother and the proband under a mixed model accounting for intra-individual correlation. The 90% credible intervals for the odds ratio of the mean in the mother and the proband to the mean in controls were 0.32–0.46 and 0.18–0.28, respectively. **l**, Day-12 differentiated megakaryocytes for the indicated individuals were stained for F-actin (red) and HDAC6 (green). Top, HDAC6 is expressed in the cytosol and is trafficked to proplatelets as shown in megakaryocytes from the control and the father (bold arrows). Middle, megakaryocytes from the proband show no HDAC6 expression while cultures from the mother contain a mixture of megakaryocytes that are positive and negative (15 of the 45 megakaryocytes) for HDAC6 expression. Bottom, only the HDAC6 staining for the proband and mother. This experiment was performed once. **m**, Day-12 differentiated megakaryocytes for the indicated individuals were stained for acetylated α-tubulin (green). Highly organized tubulin structures are present in all megakaryocytes from the control and father while the patient (47 of the 57 megakaryocytes) and mother (16 of the 46 megakaryocytes) contain megakaryocytes that show signs of tubulin depolymerization (as indicated by an asterisk). This experiment was performed once.



Extended Data Fig. 10 | Thrombocytopenia due to compound regulatory and coding rare variants in *MPL*. **a**, Top, smoothed covariance between H3K27ac ChIP–seq and ATAC-seq (as in Fig. 4a) and coverage tracks generated by RedPop for activated CD4$^+$ T cells (aCD4), B cells, erythroblasts, megakaryocytes, monocytes and resting CD4$^+$ T cells (rCD4). Middle, *MPL* gene with exons in yellow. Bottom, positions of the deletion (blue bar) and SNV (blue dot) in the proband. **b**, Pedigree for the proband with thrombocytopenia owing to a 454-bp deletion encompassing exon 10 of *MPL*, which was inherited from the mother, and an SNV just upstream of the 5′ untranslated region of *MPL*. **c**, Sanger sequencing traces confirming the presence of the heterozygous SNV in the proband and its absence in the mother. **d**, Gel electrophoresis of PCR amplicons covering the deletion confirming presence of the deletion in the proband and the mother. The PCR was conducted on two independent samples in the proband and once in the mother and the control (wt). **e**, MFI on the *y*-axis obtained by the flow cytometry measurement of MPL abundance (CD110) on the membrane of platelets from five unrelated healthy controls, the mother and the proband. The MFI was normalized to unstained platelets. We fitted a linear regression model with an intercept term representing the mean in the control, a coefficient representing the difference in means between the mother and control ($P = 0.1828$) and a coefficient representing the difference in means between the proband and control ($P = 0.0086$). Distribution summaries show mean ± s.e.m. where multiple observations are available. **f**, Results of luciferase reporter assays in K562 cells expressing empty pGL3 vector or after cloning with an MPL promoter fragment containing the wild-type G allele (MPL-SNV-G) or the variant A allele (MPL-SNV-A). The measurements were derived from $n = 4$ independent transfection experiments. The *P* values were obtained by one-way ANOVA and adjusted for multiple comparisons using Tukey's method. Distribution summaries show mean ± s.e.m.

nature research

Corresponding author(s): Ernest Turro, F Lucy Raymond, Willem H Ouwehand

Last updated by author(s): May 1, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Illumina Isaac aligner (v.SAAC00776.15.01.27); Illumina Starling variant caller (v.2.1.4.2); Illumina Manta (v.0.23.1); Illumina Canvas (v.1.1.0.5); Illumina HiSeq Analysis Software (v.2.0); BWA (v.0.7); VILMAA (https://github.com/mh11/VILMAA); CellBase (v.4.5); VEP (Ensembl API 89); OpenClinica (https://www.openclinica.com/); CiviCRM (https://civicrm.org/).
Data analysis	R (v.3.1 to v.3.5); CrossMap (v.0.2.7); samtools (v.1.3 to v.1.9); verifyBamID (v.1.1.3); bedtools (v.2.26.0); picard (v.1 to v.2); Apache Spark (v.2.5); plink (v.1.9); PRIMUS (v.1.7); Prism (v.7); RedPop (v.1; https://gitlab.haem.cam.ac.uk/et341/redpop); Blueprint DCC ChIP-Seq Analysis Pipeline; Sapientia(TM) (v.1.0 to v.1.9); IGV (v.2, v.3); F-Seq (v.1.84); deepTools plotFingerprint (v.2.3.5); MACS2 (v.2.1.1); MatInspector (https://www.swmath.org/software/21812); Genalice (http://www.genalice.com/); VILMAA (https://github.com/mh11/VILMAA) ;CellBase (v.4.5); VEP (Ensembl API 89); BWA (v.0.7); Kaluza Analysis Software (v.2.1). R packages: BeviMed, biomaRt, Biostrings, cowplot, data.table, doParallel, dplyr, egg, foreach, gdsfmt, GENESIS, GenomicRanges, GGally, ggpubr, ggplot, ggplot2, ggrepel, ggthemes, grid, gridExtra, Gviz, GWASTools, hexbin, httr, jsonlite, magrittr, MASS, Matrix, methods, ontologyIndex, parallel, plotly, plot3D, plyr, png, RColorBrewer, reshape2, SNPRelate, rtracklayer, R.utils, scales, scatterplot3d, stringr, taRifx, tibble, tidyr, VGAM, viridis, xml2, xyloplot.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Genotype and phenotype data from the 4,835 participants enrolled in the NIHR BioResource for the 100,000 Genomes Project–Rare Diseases Pilot can be accessed by seeking access via Genomics England Limited following the procedure outlined at: https://www.genomicsengland.co.uk/about-gecip/joining-research-community/. The genotype data for the 764 UK Biobank samples will be made available through a data release process which is being overseen by UK Biobank (https://www.ukbiobank.ac.uk/). The phenotype data from UK Biobank participants are available from UK Biobank using their access procedures.
Subject to ethical consent, the genotype data of the remaining 7,438 NIHR BioResource participants are available from the European Genome-phenome Archive (EGA) at the EMBL European Bioinformatics Institute under access procedures managed by EGA. The domain specific accessions are as follows: BPD: EGAD00001004519, CSVD: EGAD00001004513, EDS (EGAD00001005123), HCM: EGAD00001004514, ICP: EGAD00001004515, IRD: EGAD00001004520, LHON (EGAD00001005122), MPMT: EGAD00001004521, NDD: EGAD00001004522, NPD: EGAD00001004516, PAH: EGAD00001004525, PID: EGAD00001004523, PMG: EGAD00001004517, SMD: EGAD00001004524, SRNS: EGAD00001004518. Access to detailed phenotype data of the NIHR BioResource participants can be requested by contacting the NIHR BioResource Data Access Committee at dac@bioresource.nihr.ac.uk.

The ATAC-seq and H3K27ac ChIP-seq data to support the generation of the regulomes are available from GEO, EGA, ENCODE, or referenced to their publication. For transcription factor ChIP-seq: MK (GATA1, GATA2, TAL1, FLI1 - PMID: 21571218; MEIS1 - PMID: 25258084; CTCF - EGAD00001002362); EB (GATA1, KLF1, NFE2, TAL1 - PMID: 25521328; CTCF - EGAD00001002377); MONO (ENCSR000ATN); B (ENCSR000AUV). For H3K27ac ChIP-seq: MK (EGAD00001002362); EB (EGAD00001002377); MONO (ERR829362 (ERS257420), ERR829412 (ERS222466), ERR493634 (ERS214696), BLUEPRINT consortium); B (ERR1043004, ERR1043129, ERR928206, ERR769436, BLUEPRINT consortium); aCD4(PMID:28870212); rCD4 (PMID:28870212). For ATAC-seq: MK (EGAD00001001871); EB (SRR5489430 (GSM2594182)); MONO (EGAD00001006065); B (SRR2126769 (GSE71338)); aCD4 (GSE124867); rCD4 (GSE124867).

Reported alleles and their clinical interpretation have been deposited with ClinVar under the study names "NIHR_Bioresource_Rare_Diseases_13k", "NIHR_Bioresource_Rare_Diseases_Retinal_Dystrophy", "NIHR_Bioresource_Rare_Diseases_MYH9" and "NIHR_Bioresource_Rare_Diseases_PID".

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	13,187 samples. As our study piloted WGS of rare disease patients on a national scale, we had to accept recruitment of patients with a wide range of diseases and diverse aetiologies. Under certain realistic scenarios concerning penetrance and genetic architecture, only a small number of cases (< 10) with a shared genetic aetiology and several hundred non-cases are required to identify a genetic association. Previous WES studies with comparable sample sizes had been shown to be well powered, by replication and biological follow up.
Data exclusions	150 samples that failed quality control, as detailed in Supplementary Information. The data exclusion criteria were established over time as WGS data were generated, but were applied uniformly to the final dataset. The exclusion criteria were not informed by the phenotypes of the participants, to minimise the possibility of exclusion generating confounding.
Replication	Experimental replication was not attempted.
Randomization	For logistical reasons, recruitment and WGS were performed concurrently. Consequently, it was not possible to randomise the order of individuals to sequencing over time and, thus, over the three successive read length batches. However, we found that the variation in read length did not pose any difficulty in practice thanks to the stringent quality control imposed by our variant filters.
Blinding	Our study was not an intervention study and therefore blinding was not required. However, WGS quality control was performed without reference to the phenotypes.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	For the functional analysis of the GATA1 enhancer/HDAC6 deletion: Rabbit HDAC6 (clone D2E5, cat.no 7558S, staining 1:50, blot 1:1000, Cell Signaling technology, Danvers, MA, USA), mouse anti-acetylated tubulin antibody (clone 6-11B-1, cat. no T7451, staining 1:50, blot 1:1000, Sigma, St Louis, MO, USA), mouse anti-alpha-tubulin (clone 236-10501, cat. no A11126, staining 1:250, blot 1:1000, Thermo Fisher Scientific, Waltham, MA, USA), rabbit VWF (cat. no A0082, staining 1:50, Dako Agilent Technologies, Leuven, BE), mouse CD63 and rat GATA1 N6 (cat. nos sc-5275 and sc-265 respectively, clones MX-49.129.5 and N6 respectively, staining 1:50 -mouse- and blot 1:1000 -rat-, Santa Cruz Biotechnology, Dallas, TX, USA), rabbit GATA1 (NF that was produced against recombinant N-terminal zinc finger, blot 10 μg/ml, PMID:19924028), rabbit GAPDH (clone 14C10, cat. no 2118S, blot 1:1000, Cell Signaling) and integrin beta3 (clone H96, cat. no sc-14009, blot 1:1000, Santa Cruz Biotechnology). For MPL expression on platelets: APC-labelled IgG1 against CD42b (clone HIP1, cat. no 551061, staining 1:5, BD Pharmingen, number: 551061), PE-labelled IgG1 against CD110 (clone REA250, cat. no 130-101-648, staining 1:5, Miltenyi Biotec) and a PE-labelled isotype control (clone MOPC-21, cat.no 555749, BD Pharmingen); the staining was the same for all: add antibodies to 5ul of whole blood - make up to 12.5ul with PBS.
Validation	For the functional analysis of the GATA1 enhancer/HDAC6 deletion: The rabbit GATA1 antibody was produced against the N-terminal zinc finger of GATA1 (see PMID:19924028). This antibody was validated by immunoblot analysis against full length GATA1 expressed in HEK293 cells in parallel with the commercial GATA1 N6 antibody and both do generate bands of comparable sizes that are absent from lysates of non-transfected cells. All the other antibodies used have been published by others as specified in the datasheets from the suppliers mentioned above. For the MPL expression on platelet: all antibodies were used according to manufacturer's instruction.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Age: birth to 95 years old. Gender: Male and Female. Patients with rare disorders across 15 disease domains, and relatives. Wide range of diagnosis and treatment categories, as detailed in Supplementary Information.
Recruitment	Patients were recruited from 83 hospitals in the UK and worldwide, as detailed in Supplementary Information. The patient populations at these hospitals differ with respect to genetic ancestry, which may have induced a degree of selection bias. This potential bias was mitigated by enrolling as widely as possible across different hospitals (see Extended Data Figure 1a) and by accounting for coarse ancestry in the association analyses.
Ethics oversight	East of England Cambridge South national research ethics committee (REC) reference 13/EE/0325 or separate local ethics, as detailed in Supplementary Information.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	No ChIP-seq data were generated, we used publicly available data. Transcription factor ChIP-seq: MK - GATA1, GATA2, TAL1, and FLI1: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE24674; MK - MEIS1: https://www.ebi.ac.uk/ega/datasets/EGAD00001000745; MK - CTCF: https://www.ebi.ac.uk/ega/datasets/EGAD00001002362; EB - GATA1, KLF1, NFE2 and TAL1: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE59801; EB - CTCF: https://www.ebi.ac.uk/ega/datasets/EGAD00001002377; MONO - CTCF: https://www.encodeproject.org/experiments/ENCSR000ATN; B - CTCF: https://www.encodeproject.org/experiments/ENCSR000AUV. H3K27ac ChIP-seq:

	MK: https://www.ebi.ac.uk/ega/datasets/EGAD00001002362; EB: https://www.ebi.ac.uk/ega/datasets/EGAD00001002377; MONO: BLUEPRINT Consortium website (http://www.blueprint-epigenome.eu) with accession IDs ERR829362 (ERS257420), ERR829412 (ERS222466), ERR493634 (ERS214696); B: BLUEPRINT Consortium website (http://www.blueprint-epigenome.eu) with accession IDs ERR1043004, ERR1043129, ERR928206, ERR769436; aCD4: https://www.ebi.ac.uk/ega/datasets/EGAD00001002686; rCD4: https://www.ebi.ac.uk/ega/datasets/EGAD00001002686.
Files in database submission	n/a
Genome browser session (e.g. UCSC)	n/a

Methodology

Replicates	As in publications (MK - PMID:25258084 and PMID:21571218; EB - PMID:25521328; aCD4 and rCD4 - PMID:28870212) or the ENCODE website (MONO - https://www.encodeproject.org/experiments/ENCSR000ATN/; B - https://www.encodeproject.org/experiments/ENCSR000AUV).
Sequencing depth	As in publications (MK - PMID:25258084 and PMID:21571218; EB - PMID:25521328; aCD4 and rCD4 - PMID:28870212) or the ENCODE website (MONO - https://www.encodeproject.org/experiments/ENCSR000ATN/; B - https://www.encodeproject.org/experiments/ENCSR000AUV).
Antibodies	As in publications (MK - PMID:25258084 and PMID:21571218; EB - PMID:25521328; aCD4 and rCD4 - PMID:28870212) or the ENCODE website (MONO - https://www.encodeproject.org/experiments/ENCSR000ATN/; B - https://www.encodeproject.org/experiments/ENCSR000AUV).
Peak calling parameters	TFs and H3K27ac peaks were called with MACS2, significance threshold was set to qvalue<1e-5, narrow option was used.
Data quality	Low quality reads (-q 15), multi-mapped and duplicate reads were marked and removed with samtools and picard respectively. ChIP-seq efficiency was assessed with deepTools fingerPrint.
Software	BWA (v.0.7); picard (v.1 to v.2); deepTools plotFingerprint (v.2.3.5); MACS2 (v.2.1.1)

Flow Cytometry

Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	For MPL expression on platelets: the level of MPL protein on the platelet membrane was measured by flow cytometry (Beckman Coulter FC500) using the monoclonal antibodies: APC-labelled IgG1 against CD42b (clone HIP1, BD Pharmingen, cat. no 551061), PE-labelled IgG1 against CD110 (clone REA250, Miltenyi Biotec, cat. no 130-101-648) and a PE-labelled isotype control (clone MOPC-21, BD Pharmingen, cat. no 555749). In short, a sample of EDTA anticoagulated blood was incubated with anti-CD110 (or control) and anti-CD42b for 30 minutes.
Instrument	Beckman Coulter FC500
Software	Kaluza Analysis Software from Beckman (Version 2.1)
Cell population abundance	n/a
Gating strategy	Platelets were gated based on size using forward scatter and size scatter. The median fluorescent intensity of the CD110 PE-antibody was calculated for all CD42b positive platelets.

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

A map of object space in primate inferotemporal cortex

https://doi.org/10.1038/s41586-020-2350-5

Received: 21 January 2019

Accepted: 17 March 2020

Pinglei Bao[1,2]✉, Liang She[1], Mason McGill[3] & Doris Y. Tsao[1,2,3]✉

The inferotemporal (IT) cortex is responsible for object recognition, but it is unclear how the representation of visual objects is organized in this part of the brain. Areas that are selective for categories such as faces, bodies, and scenes have been found[1–5], but large parts of IT cortex lack any known specialization, raising the question of what general principle governs IT organization. Here we used functional MRI, microstimulation, electrophysiology, and deep networks to investigate the organization of macaque IT cortex. We built a low-dimensional object space to describe general objects using a feedforward deep neural network trained on object classification[6]. Responses of IT cells to a large set of objects revealed that single IT cells project incoming objects onto specific axes of this space. Anatomically, cells were clustered into four networks according to the first two components of their preferred axes, forming a map of object space. This map was repeated across three hierarchical stages of increasing view invariance, and cells that comprised these maps collectively harboured sufficient coding capacity to approximately reconstruct objects. These results provide a unified picture of IT organization in which category-selective regions are part of a coarse map of object space whose dimensions can be extracted from a deep network.

Object recognition, the process by which distinct visual forms are assigned distinct identity labels, lies at the heart of our ability to make sense of the visual world. It underlies many neural processes that operate on objects, including consciousness, attention, visual memory, decision making, and language. Befitting the central importance and computational complexity of object recognition, a large volume of the brain, IT cortex, is dedicated to solving this challenge[7].

One of the most striking features of IT is the existence of several distinct anatomical networks that are specialized for processing specific categories[2,4,5] or stimulus dimensions[8–11]. However, these networks comprise only part of IT, and much of IT is not differentially activated by any known stimulus comparison. Here we investigate whether this 'unexplained' IT shows any functional specialization. Furthermore, beyond simply parcelling IT, we investigate whether there is an overarching general principle governing the anatomical layout of IT cortex.

Many previous studies have tried to address this latter question, but the answers obtained remain piecemeal. Early studies using electrophysiology in monkeys suggested a columnar architecture for visual shape[12], but the small field-of-view of electrophysiology precluded understanding the larger-scale organization of these columns. Later studies, using functional MRI (fMRI) in humans, proposed various schemes to explain large-scale IT organization including retinotopy[13] and real-world size[14], but these proposals did not provide a complete account of IT organization and lacked ground-truth validation at the level of single units. Here, we combined fMRI, electrical microstimulation, and electrophysiology in the same animals to investigate the organization of macaque IT at multiple scales, and found that a large portion of macaque IT cortex is topographically organized into a map of object space that is repeated three times.

Identifying a new IT network

To discover the functional specialization of still unexplained parts of IT cortex, one strategy would be to guess. However, lacking any good guesses, we decided to approach the problem from an anatomical perspective. We ran a large set of stimulus comparisons to localize face, body, scene, colour, and disparity patches in a specific monkey (M1) and thereby define the 'no man's land' of IT cortex in this monkey: regions that were not identified by any known localizer (Fig. 1a, b). We then electrically microstimulated a random site within this no man's land in central IT cortex[15]. This experiment revealed that the stimulated region (NML2) was connected to two other, discrete regions in IT (NML1, NML3) (Fig. 1b, Extended Data Fig. 1), forming a previously unknown anatomical network within no man's land.

To understand the function of this new network, we first recorded the neural responses of cells in the three patches to 1,224 images, consisting of 51 objects each presented at 24 views belonging to 6 different categories (Extended Data Fig. 2a, b). Responses were remarkably consistent (Fig. 2a, Extended Data Fig. 3a–d). Cells in all three patches responded minimally to faces. Their preferred stimuli, while consistent across patches, were not confined to any one semantic category (Fig. 2a).

To investigate whether this network exists in every animal, we identified the five most- and least-preferred objects of the network based

---

[1]Division of Biology and Biological Engineering, Tianqiao and Chrissy Chen Institute for Neuroscience, Caltech, Pasadena, CA, USA. [2]Howard Hughes Medical Institute, Caltech, Pasadena, CA, USA. [3]Computation and Neural Systems, Caltech, Pasadena, CA, USA. ✉e-mail: pbao@caltech.edu; dortsao@caltech.edu

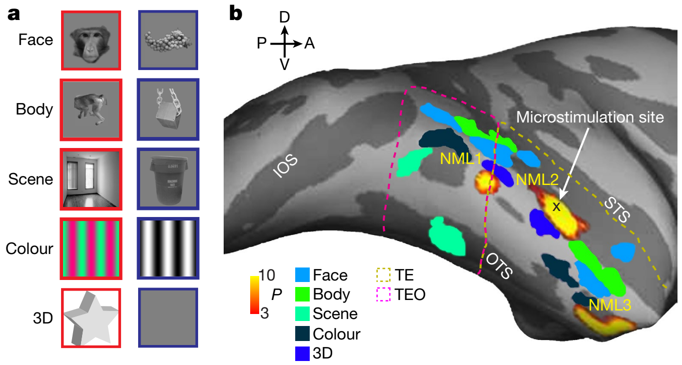


Fig. 1 | Microstimulation reveals a new anatomical network in IT cortex. **a**, Stimulus contrasts used to identify known networks in IT (see Methods). **b**, Inflated brain (right hemisphere) for monkey M1 showing known IT networks mapped in this animal. Regions activated by microstimulation of NML2 are shown in yellow. All activation maps shown at a threshold of $P < 10^{-3}$, not corrected for multiple comparisons. Yellow and magenta outlines indicate the boundaries of TE and TEO, respectively[31].

on mean responses of cells recorded from monkey M1 (Fig. 2a). We presented these stimuli to monkey M1 in an fMRI experiment and confirmed that the resulting map overlapped that revealed by microstimulation (Fig. 2e). We then presented these stimuli to three other monkeys (M2–M4) and found similar networks in all three animals (Fig. 2e). Single-unit recordings targeted to this network in monkey M2 revealed a response pattern that was highly consistent with that in monkey M1 (Fig. 2a) (Pearson correlation of the mean responses to each object between monkeys M1 and M2, $r = 0.89$, $P < 10^{-16}$). This justifies referring to an 'NML network' across animals.

In the face patch network, neurons in posterior patches are view-specific whereas those in the most anterior patch are view-invariant[16]. We found a similar difference between the three NML patches in terms of their view invariance. Significantly more cells in NML3 were view-invariant than in NML1 (two-tailed *t*-test; $t(137) = 5.10$, $P < 10^{-5}$; Extended Data Fig. 3e). Population similarity matrices to objects at different views also showed an increase in view invariance going anteriorly, with emergence of parallel diagonal stripes in the NML3 similarity matrix (Fig. 3a (top), Extended Data Fig. 3f). Notably, many cells showed view invariance to objects that the monkey had not experienced, such as an aeroplane (Fig. 3b (top)).

Next, we investigated what is being coded by cells in this network. Scrutinizing the most- and least-preferred objects (Fig. 2a (bottom)), we noticed that all of the preferred objects contained thin protrusions, whereas the non-preferred objects were round. This suggested that one feature NML neurons might be selective for is high aspect ratio. We confirmed this using both responses to the original object image set (Extended Data Fig. 3g, see Methods) as well as a simplified stimulus set consisting of a line segment independently varied in aspect ratio, curvature, and orientation (Fig. 2f, Extended Data Fig. 2c). Thus a common preferred feature of cells in the NML network is high aspect ratio.

NML cells encode axes of object space

We next attempted to identify the relevant shape dimensions for the NML network in a systematic way that does not depend on subjective visual inspection. Until recently, this was difficult because of the lack of a computational scheme to parametrize arbitrary objects. Deep networks trained to classify objects provide a powerful solution to this problem[17]. They allow parametrization of arbitrary objects through computation of a few thousand numbers, the unit activations in a deep layer. To make the parametrization even more compact, one can perform principal components analysis (PCA) on these unit activations. We built an object space by passing the stimulus set we presented to the monkey (Extended Data Fig. 2a, b) through AlexNet, a deep network trained on object classification[6], and then performing PCA on the responses of units in layer fc6 of this network (Extended Data Fig. 4a). The first principal component (PC) corresponds roughly to things with protrusions (spiky) versus those without (stubby) (Extended Data Fig. 4b). The second PC corresponds roughly to animate versus inanimate (note that we use 'animate' and 'inanimate' as shape descriptors without any semantic connotation). We determined that 50 object dimensions could explain 85% variance in the AlexNet fc6 response (Extended Data Fig. 4c) and thus used 50 dimensions in the remaining analyses. We then analysed the responses of cells in the NML network by computing a 'preferred axis' for each cell through linear regression, namely, the coefficients $\mathbf{c}$ in the equation $R = \mathbf{c}\cdot\mathbf{f} + c_0$, where R is the response of the cell, $\mathbf{f}$ is the 50D object feature vector, and c_0 is a constant offset (see Methods).

Cells showed significant tuning to many of the 50 object dimensions (Pearson correlation $P < 10^{-3}$ between feature values and neural responses). On average, each cell was significantly tuned to seven dimensions. Notably, the preferred axis of each cell was stable to the precise image set (Extended Data Fig. 5a). The 50D linear object space model could explain 44.7% variance, or 53.3% of the explainable variance of NML neurons on average (Extended Data Fig. 5b); this is significantly higher than a Gaussian model and similar to a quadratic model (Extended Data Fig. 5c, d). Consistent with the high explained variance by the linear model, cell tuning along the preferred axis in the 50D object space was ramp-shaped (Fig. 3c, top). Similar ramp-shaped tuning has previously been reported for face-selective cells[18]. NML neurons also showed approximately flat tuning along orthogonal axes (Extended Data Fig. 5e), another property that has been previously observed in face-selective cells[18]. Together, ramp-shaped tuning along the preferred axis and flat tuning along orthogonal axes implies that cells in the NML network are linearly projecting incoming objects, formatted as vectors in object space, onto specific preferred axes.

Overall, the organization and code of the NML network are strikingly similar to those of the face patch network. The NML network consists of connected patches, cells within the network show a consistent pattern of selectivity, there is increasing view invariance along the network, and finally, single cells in the network represent object identity through an axis code. Thus there seems to be a clear structural parallel between the face network and the NML network. We therefore investigated whether additional networks in IT cortex follow the same scheme.

The body network follows the same scheme

We next recorded from the macaque body network, a set of regions adjacent to face patches that respond more to animate compared to inanimate objects[4] (Fig. 2b), as well as the face network (Fig. 2c). Population similarity matrices showed increased view invariance in the most anterior body patch (Fig. 3a, b (middle), Extended Data Fig. 3e, f), consistent with a previous study[19]. Cells in the body network also showed ramp-shaped tuning along their preferred axes (Fig. 3c (middle), Extended Data Fig. 5a) and flat tuning along orthogonal axes (Extended Data Fig. 5e). Thus the body network follows the same general anatomical organization and coding scheme as the NML and face networks.

A general rule governing IT organization

The finding of three networks (NML, body and face) that all follow the same organization and coding scheme suggests that there might be a general principle that governs the organization of IT cortex. Recall that the first two axes of object space are roughly stubby versus spiky, and animate versus inanimate (Extended Data Fig. 4b). We noticed a remarkable relationship between these two axes and the selectivity of

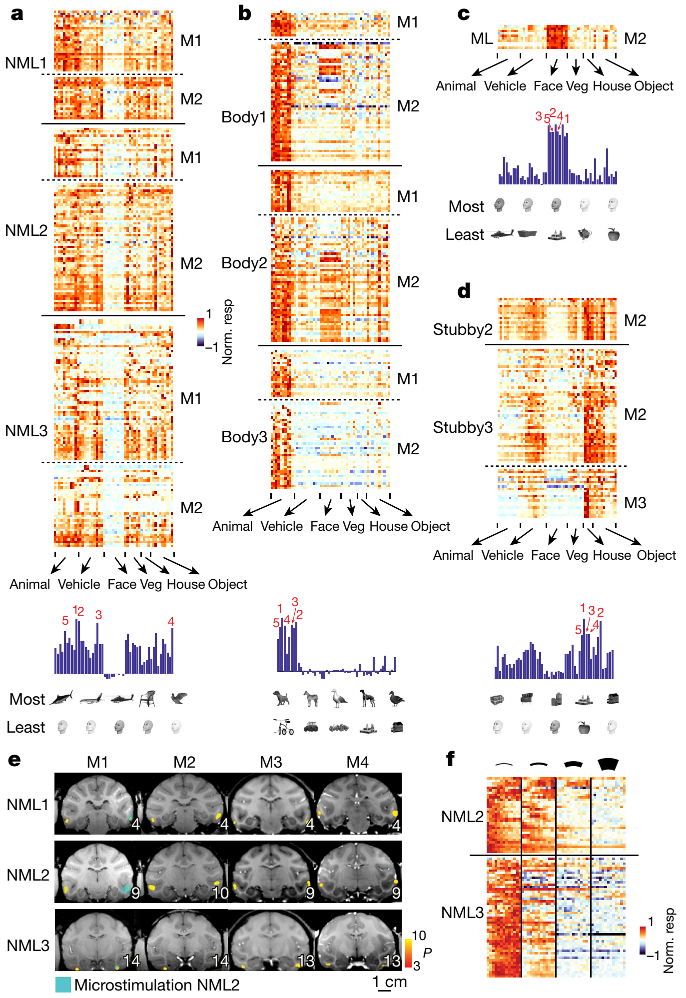


Fig. 2 | Distinct object preferences among four different networks in IT cortex. a–d, Top, responses of cells to 51 objects from six different categories. Responses to each object were averaged across 24 views. Cells were recorded in three patches (NML1, NML2 and NML3) from the NML network (**a**); in three patches of the body network (**b**); in patch ML of the face network (**c**); and in two patches of the stubby network (**d**). Middle, blue charts show average responses to each object in each network. Numbers indicate the five most-preferred objects. Bottom, five most-preferred (top row) and least-preferred (bottom row) objects for each network, based on averaged responses; images 1 to 5 are shown from left to right. **e**, Coronal slices containing NML1, NML2, and NML3 from monkeys M1, M2, M3, and M4 showing difference in activation in response to the five most-preferred versus five least-preferred objects determined from electrophysiology in the NML network of monkey M1. In M1, the microsimulation result is also shown as a cyan overlay with threshold $P < 10^{-3}$, uncorrected. Inset numbers indicate AP coordinate relative to interaural 0[31]. **f**, Responses of cells from patches NML2 and NML3 of the NML network to a line segment that varied in aspect ratio, curvature, and orientation. Responses are averaged across orientation, and curvature runs from low to high from left to right for each aspect ratio. Aspect ratio accounts for 22.8% of response variance on average across cells, curvature for 5.6% of variance, and orientation for 3.5% of variance.

the NML, body, and face networks. Face patches prefer stubby animate objects; body patches prefer spiky, animate objects; and NML patches prefer spiky objects regardless of animacy (Fig. 2a). These observations made us wonder whether all of IT might be topographically organized according to the first two dimensions of object space (Fig. 4a), in the same way that retinotopic cortex is organized according to polar angle and eccentricity.

As a first step to test this hypothesis, we projected all the stimuli that we showed to the monkey onto the first two dimensions of object space, and marked the top 100 images for the NML, body, and face networks (Fig. 4b; orange, green, and blue dots). They approximately spanned three quadrants of the space. If IT cortex is indeed laid out according to the first two dimensions of object space, we predicted there should be a fourth network representing objects that project strongly onto the remaining unrepresented quadrant—namely stubby, inanimate objects without protrusions (for example, a USB stick or radio).

To test this prediction, we first ran an fMRI experiment with four blocks, corresponding to the four quadrants of object space (Fig. 4a). Comparison of stubby versus other blocks revealed a network that contained multiple patches selective for stubby objects (Fig. 4c). Electrophysiology targeted to two of these patches revealed cells that were strongly selective for stubby objects (Fig. 2d), whose preferred axes occupied the previously unrepresented quadrant (Fig. 4b, magenta dots). The general properties of the stubby network were very similar to those of the NML, face, and body networks. Population similarity matrices showed increased view invariance in the most anterior stubby patch (Fig. 3a, b (bottom), Extended Data Fig. 3f). Cells in the stubby network also showed ramp-shaped tuning along their preferred axes (Fig. 3c (bottom), Extended Data Fig. 5a) and flat tuning along orthogonal axes (Extended Data Fig. 5e). Thus, the hypothesis that IT is organized according to the first two dimensions of object space revealed a second new shape network.

One potential concern is that the 51 objects at 24 views that we used to assess the selectivity of cells in each network were too sparse and may not have allowed identification of the true selectivity of cells. We presented 1,593 completely different objects to a subset of cells in the NML, body, and stubby networks and found responses consistent with those to our original stimulus set (Extended Data Fig. 6a, b). In particular, preferred axes measured using the new stimuli segregated into three different regions of object PC1–PC2 space (Extended Data Fig. 6a), and the preferred stimuli of each network were qualitatively similar to those identified using the original stimuli (Extended Data Fig. 6b).

It might seem suspiciously serendipitous for IT to be organized according to the first two dimensions of an object space computed using a specific image set with a specific deep convolutional network. In fact, these first two axes do not depend strongly on the particular image set (Extended Data Fig. 4d–f) or network (Extended Data Fig. 4g–j) used to compute them (see Supplementary Information).

A map of object space

What is the anatomical layout of the face, body, NML, and stubby networks? An overlay of the four networks onto coronal slices and a cortical flat map revealed a remarkably ordered progression (Fig. 4c, d; see Extended Data Fig. 7 for response time courses from each patch). There is a clear sequence from body to face to stubby to NML in both hemispheres that is repeated in the same order in posterior, middle, and anterior IT. This pattern was consistent across animals (Fig. 4c, d) and confirmed by quantitative analysis of the linear fit between patch-ordered label and cortical location of patch peak ($P < 10^{-18}$ for posterior, middle, and anterior IT, Fig. 4e–g). This strikingly regular progression suggests the existence of a coarse map of object space that is repeated at least three times, with increasing view invariance at each stage.

These four networks, together with the disparity, scene, and colour networks, occupy about 53% of IT cortex, so additional networks may exist. Not all of the networks consisted of exactly three patches; for example, the stubby and NML networks each contained four patches (Fig. 4d, see Supplementary Information), and previous work has suggested that there are six face patches in each hemisphere, with some individual variability[20]. Thus, IT cortex may contain additional repetitions of the object space map. Furthermore, we emphasize that our study addresses IT organization at a coarse spatial scale and does not exclude the possibility of additional organization at finer spatial

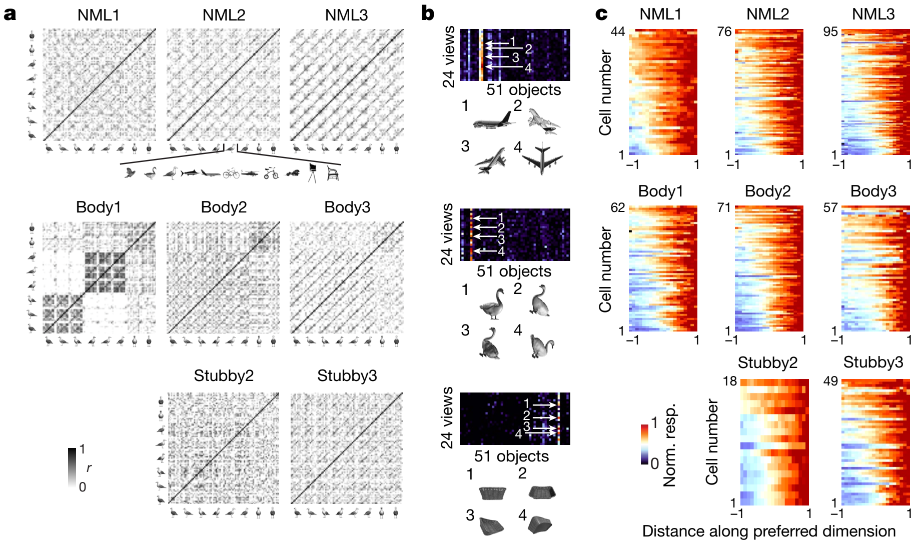


Fig. 3 | Each network contains a hierarchy of increasingly view-invariant nodes, and single cells in each node show ramp-shaped tuning. **a**, Population similarity matrices in the three patches of the NML network (top), three patches of the body network (middle) and two patches of the stubby network (bottom) pooled across monkeys M1 and M2. An 88 × 88 matrix of correlation coefficients was computed from responses of cells in each patch to 88 stimuli (8 views × top 11 preferred objects). **b**, Responses from three example cells recorded in NML3 (top), the body network (middle) and the stubby network (bottom) to 51 objects at 24 views. Four different views of the most preferred object are shown below each response matrix. **c**, Responses of neurons recorded from patches in the NML network (top), the body network (middle) and the stubby network (bottom) as a function of distance along the preferred axis. The abscissa is rescaled so that the range [−1,1] covers 95% of the stimuli. Half the stimulus trials were used to compute the preferred axis for each cell, and held-out data were used to plot the responses shown.

scales (Extended Data Fig. 8; see Supplementary Information). Recordings from multiple grid holes suggest that each patch spans 3–4 mm (Extended Data Fig. 8a–d). Although we failed to find clustering at finer scales within a patch (Extended Data Fig. 8e, f) or clustering for any dimensions beyond the first two (Extended Data Fig. 8g, h), it is possible that mapping techniques with higher spatial resolution may reveal additional substructure within patches.

If the first two dimensions of object space derived from a deep network are indeed meaningful in terms of brain representation, we should be able to design novel stimuli to identify the four networks. To this end, we generated three new image sets (silhouettes, fake objects, and deep dream images) with very different properties from those of the original image set of Fig. 4a. In each case, fMRI revealed four networks similar to those in Fig. 4c (Extended Data Fig. 6c–e).

Explaining previous accounts of IT

The principle that IT cortex is organized according to the first two axes of object space provides a unified explanation for many previous observations concerning the functional organization of IT, including not only the existence of face[1] and body areas[3], but also gradients for representing animate versus inanimate and small versus large objects[14] (Extended Data Fig. 9a, b), a gradient for representing open versus closed topologies[21] (Extended Data Fig. 9c), the curvature network[11] (Extended Data Fig. 9d), and the visual word form area[22] (Extended Data Fig. 9e). Furthermore, within category-selective regions, the object space model explains activity better than the semantic category hypothesis[23] (Extended Data Fig. 10). Overall, these results demonstrate the large explanatory power of the object space model.

Reconstructing general objects

We next investigated the richness of the feature space represented by cells in the four networks that comprise the map of object space. To quantify the object information available in the map of object space formed by the four networks, we attempted to decode object identity using the responses of cells from these networks. We used leave-one-object-out cross-validation to learn the linear transform that maps responses to features (Extended Data Fig. 11a, b). The explained variance for each dimension showed that many dimensions are coded in each network beyond the first two (Extended Data Fig. 11c), allowing a target object to be identified among distractors (Extended Data Fig. 11d–f).

To directly visualize the information about object features that is carried by neurons in these four networks, we attempted to reconstruct general objects using neural activity. We passed decoded object feature vectors through a generative adversarial network trained to invert layer fc6 of AlexNet[24]. Reconstructions were impressively accurate in details (Fig. 5a). Figure 5b shows the distribution of normalized reconstruction distances between the actual and best possible reconstructions (see Methods). As a second method to recover objects from neural activity, we searched a large auxiliary object database for the object with a feature vector closest to that decoded from neural activity. This method also yielded recovered images that picked up many fine structural details (Extended Data Fig. 11g). Overall, these results suggest that the four networks of the IT object space map are sufficient to encode a reasonably complete representation of general objects, and thus the number of networks used to solve general object recognition need not be astronomically high.

Discussion

We have shown that IT contains a coarse map of object space that is repeated three times, with increasing invariance at each stage. This map consists of at least four regions that tile object space. This map parsimoniously accounts for the previously reported face and body networks, as well as two new networks: the NML network and the stubby network. Single cells in each of the four networks use a coding principle similar to that previously identified for the face network—projection of incoming objects, formatted as points in object space, onto a preferred axis. The four networks that comprise the IT object-topic map, together with the scene, colour, and disparity networks, cover about

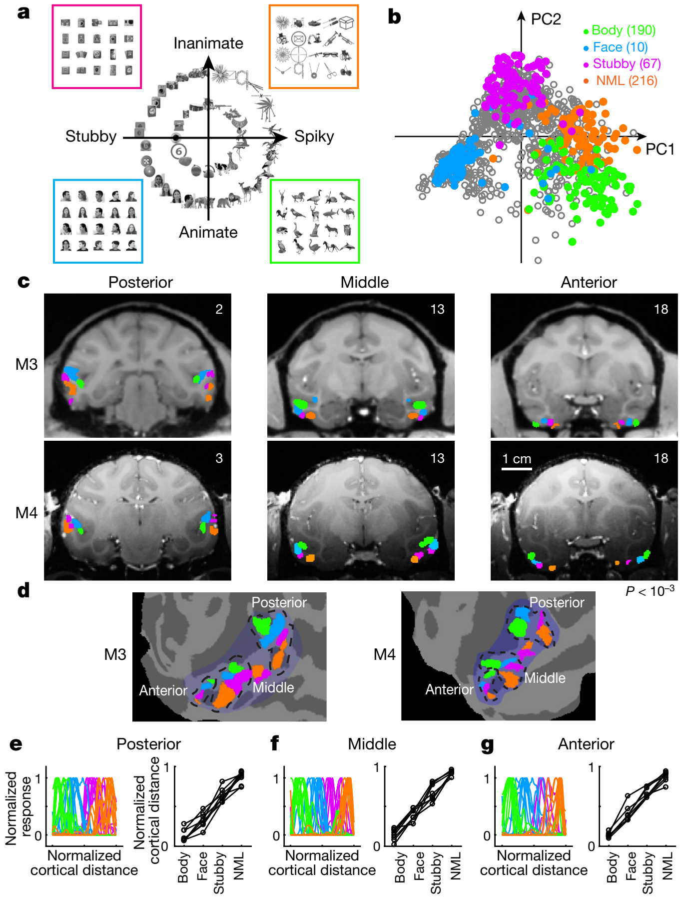


Fig. 4 | A map of object space revealed by fMRI. a, A schematic plot showing the map of objects generated by the first two PCs of object space. The stimuli in the rectangular boxes were used for mapping the four networks shown in **c**, **d** using fMRI. **b**, All the stimuli used in the electrophysiology experiments (Extended Data Fig. 2a, b) projected onto the first two dimensions of object space (grey circles). For each network, the top 100 preferred images are marked (body network: green, face network: blue, stubby network: magenta, NML network: orange). Numbers in parentheses indicate the number of neurons recorded from each network. **c**, Coronal slices from posterior, middle, and anterior IT of monkeys M3 and M4 showing the spatial arrangement of the four networks (maps thresholded at $P < 10^{-3}$, uncorrected). Here, the networks were computed using responses to the stimuli in **a**. **d**, As in **c**, showing the four networks in monkeys M3 and M4 overlaid on a flat map of the left hemisphere. **e**, Left, spatial profiles of the four patches along the cortical surface within posterior IT for data from two hemispheres of four animals. The *y*-axis shows the normalized significance level for each comparison of each voxel, and the *x*-axis shows the position of the voxel on the cortex (see Methods). Right, anatomical locations of the peak responses plotted against the sequence of quadrants in object space. **f**, **g**, As in **e** for voxels from middle IT (**f**) and anterior IT (**g**).

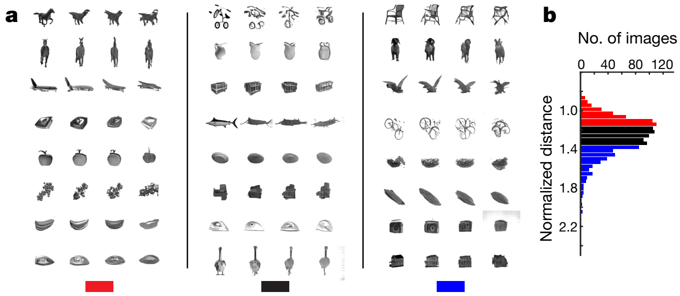


Fig. 5 | Reconstructing objects using neuronal responses from the IT object-topic map. a, Reconstructions using 482 cells from NML, body, stubby, and face networks. Example reconstructed images from the three groups defined in **b** are shown. Each row (group) of four images shows from left to right: 1, the original image; 2, the reconstructed image using the fc6 response pattern to the original image; 3, the reconstructed image using the fc6 response pattern projected onto the 50D object space; and 4, the reconstructed image based on neuronal data. **b**, Distribution of normalized distances between reconstructed feature vectors and best-possible reconstructed feature vectors (see Methods).

53% of IT. Pooling responses across the four networks enabled reasonable reconstruction of general objects, suggesting that these four networks provide a basis that spans general object space. By showing that the modular organization previously thought to be unique to a few categories may actually extend across a much larger swath of IT, we provide a powerful new map for experiments that require spatially specific interrogation of object representations.

It remains unknown whether borders between the patches are continuous or discrete[25], as fMRI-guided single-unit recording is not ideal for mapping sub-millimetre-scale structure. If the borders turn out to be continuous, this would imply that the entire notion of IT modularity may be an artefact of limited field of view. On the other hand, if the borders turn out to be discrete, this would suggest that additional factors (for example, extensive experience with specific categories[26]) may support the formation of uniquely specialized modules of cortex. The coarse map of object space identified here provides a foundation for future fine-scale mapping studies to tackle this question.

The finding that neurons in IT are clustered according to axis similarity resonates with recent approaches to unsupervised learning of object representations that seek optimal clustering of data in low-dimensional embeddings[27]. It will be important to understand why IT physically clusters neurons with similar axes—something not currently implemented in deep networks. One possible reason is that physical clustering may help to refine object representations through lateral inhibition and aid object identification in clutter[28].

Our results cast the face patch system in a new light. Previously, it was thought that the face system, with its striking clustering of face-selective cells, was a unique evolutionary consequence of the importance of face recognition to primate social behaviour. Here we show that the face system arises naturally from the statistical structure of object space. One prediction is that face-deprived animals should still show a network specialized for round objects (for example, clocks, apples), even if it is not specialized for faces per se. Selectivity for additional features may develop with face experience[26].

Our hypothesis that IT cortex is organized according to the first two dimensions of object space makes multiple new predictions. We have already confirmed several of these, including the existence of the stubby network (see Supplementary Information). Additional new predictions are that lesions in any part of IT should lead to agnosias in specific sectors of object space[29], and that other brain regions that contain face patches may also harbour maps of object space[30]. Finally, it will be important to discover whether remaining unaccounted-for regions of IT can be explained within the same general framework of a map of object space.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2350-5.

1. Kanwisher, N., McDermott, J. & Chun, M. M. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* **17**, 4302–4311 (1997).

2. Tsao, D. Y., Freiwald, W. A., Knutsen, T. A., Mandeville, J. B. & Tootell, R. B. Faces and objects in macaque cerebral cortex. *Nat. Neurosci*. **6**, 989–995 (2003).
3. Downing, P. E., Jiang, Y., Shuman, M. & Kanwisher, N. A cortical area selective for visual processing of the human body. *Science* **293**, 2470–2473 (2001).
4. Popivanov, I. D., Jastorff, J., Vanduffel, W. & Vogels, R. Heterogeneous single-unit selectivity in an fMRI-defined body-selective patch. *J. Neurosci*. **34**, 95–111 (2014).
5. Kornblith, S., Cheng, X., Ohayon, S. & Tsao, D. Y. A network for scene processing in the macaque temporal lobe. *Neuron* **79**, 766–781 (2013).
6. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet classification with deep convolutional neural networks. *Adv. Neural Information Proc. Syst*. **25**, 1097–1105 (2012).
7. Gross, C. G., Rocha-Miranda, C. E. & Bender, D. B. Visual properties of neurons in inferotemporal cortex of the macaque. *J. Neurophysiol*. **35**, 96–111 (1972).
8. Lafer-Sousa, R. & Conway, B. R. Parallel, multi-stage processing of colors, faces and shapes in macaque inferior temporal cortex. *Nat. Neurosci*. **16**, 1870–1878 (2013).
9. Verhoef, B.-E., Bohon, K. S. & Conway, B. R. Functional architecture for disparity in macaque inferior temporal cortex and its relationship to the architecture for faces, color, scenes, and visual field. *J. Neurosci*. **35**, 6952–6968 (2015).
10. Janssen, P., Vogels, R. & Orban, G. A. Selectivity for 3D shape that reveals distinct areas within macaque inferior temporal cortex. *Science* **288**, 2054–2056 (2000).
11. Yue, X., Pourladian, I. S., Tootell, R. B. & Ungerleider, L. G. Curvature-processing network in macaque visual cortex. *Proc. Natl Acad. Sci. USA* **111**, E3467–E3475 (2014).
12. Fujita, I., Tanaka, K., Ito, M. & Cheng, K. Columns for visual features of objects in monkey inferotemporal cortex. *Nature* **360**, 343–346 (1992).
13. Levy, I., Hasson, U., Avidan, G., Hendler, T. & Malach, R. Center-periphery organization of human object areas. *Nat. Neurosci*. **4**, 533–539 (2001).
14. Konkle, T. & Oliva, A. A real-world size organization of object responses in occipitotemporal cortex. *Neuron* **74**, 1114–1124 (2012).
15. Moeller, S., Freiwald, W. A. & Tsao, D. Y. Patches with links: a unified system for processing faces in the macaque temporal lobe. *Science* **320**, 1355–1359 (2008).
16. Freiwald, W. A. & Tsao, D. Y. Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science* **330**, 845–851 (2010).
17. Yamins, D. L. et al. Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proc. Natl Acad. Sci. USA* **111**, 8619–8624 (2014).
18. Chang, L. & Tsao, D. Y. The code for facial identity in the primate brain. *Cell* **169**, 1013–1028.e1014 (2017).
19. Kumar, S., Popivanov, I. D. & Vogels, R. Transformation of visual representations across ventral stream body-selective patches. *Cereb. Cortex* **29**, 215–229 (2019).
20. Tsao, D. Y., Moeller, S. & Freiwald, W. A. Comparing face patch systems in macaques and humans. *Proc. Natl Acad. Sci. USA* **105**, 19514–19519 (2008).
21. Vaziri, S., Carlson, E. T., Wang, Z. & Connor, C. E. A channel for 3D environmental shape in anterior inferotemporal cortex. *Neuron* **84**, 55–62 (2014).
22. McCandliss, B. D., Cohen, L. & Dehaene, S. The visual word form area: expertise for reading in the fusiform gyrus. *Trends Cogn. Sci*. **7**, 293–299 (2003).
23. Baldassi, C. et al. Shape similarity, better than semantic membership, accounts for the structure of visual object representations in a population of monkey inferotemporal neurons. *PLoS Comput. Biol*. **9**, e1003167 (2013).
24. Dosovitskiy, A. & Brox, T. Generating images with perceptual similarity metrics based on deep networks. *Adv. Neural Information Proc. Syst*. **29**, 658–666 (2016).
25. Aparicio, P. L., Issa, E. B. & DiCarlo, J. J. Neurophysiological organization of the middle face patch in macaque inferior temporal cortex. *J. Neurosci*. **36**, 12729–12745 (2016).
26. Arcaro, M. J., Schade, P. F., Vincent, J. L., Ponce, C. R. & Livingstone, M. S. Seeing faces is necessary for face-domain formation. *Nat. Neurosci*. **20**, 1404–1412 (2017).
27. Zhuang, C., Zhai, A. L. & Yamins, D. in *Proc. IEEE Intl Conf. Computer Vision* 6002–6012 (2019).
28. Bao, P. & Tsao, D. Y. Representation of multiple objects in macaque category-selective areas. *Nat. Commun*. **9**, 1774 (2018).
29. Rajalingham, R. & DiCarlo, J. J. Reversible inactivation of different millimeter-scale regions of primate IT results in different patterns of core object recognition deficits. *Neuron* **102**, 493–505.e495 (2019).
30. Haile, T. M., Bohon, K. S., Romero, M. C. & Conway, B. R. Visual stimulus-driven functional organization of macaque prefrontal cortex. *Neuroimage* **188**, 427–444 (2019).
31. Saleem, K. S. & Logothetis, N. K. *A Combined MRI and Histology Atlas of the Rhesus Monkey Brain in Stereotaxic Coordinates* (Academic, 2012).

Methods

Five male rhesus macaques (*Macaca mulatta*) between 5 and 8 years old were used in this study. All procedures conformed to local and US National Institutes of Health guidelines, including the US National Institutes of Health Guide for Care and Use of Laboratory Animals. All experiments were performed with the approval of the Caltech Institutional Animal Care and Use Committee.

No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment.

Visual stimuli

Stimuli for electrophysiology experiments. Three different stimulus sets were used. 1) A set of 51 objects from 6 different categories, each presented at 24 different views (Extended Data Fig. 2a, b). Except for face models, other 3D models were downloaded from https://www.3d66.com. Face 3D models were generated by Facegen (Singular Inversions) software using random parameters. The images at 24 views for each object were generated using 3dMax (Autodesk) software. Each image was presented for 250 ms interleaved with 150 ms of a grey screen. Each image was presented 4–8 times. 2) A set of line segments that varied along three dimensions: curvature, aspect ratio, and orientation (Extended Data Fig. 2c). Each image was presented for 150 ms interleaved with 150 ms of a grey screen. Each image was presented 6–8 times. 3) A set of object images consisting of 1,392 different images downloaded from www.freepngs.com. We also included 201 face images from the FEI database (https://fei.edu.br/~cet/facedatabase.html). Thus there were 1,593 images in total (Extended Data Fig. 2d). Each image was presented for 150 ms interleaved with 150 ms of a grey screen. Each image was presented 4–8 times.

Localizer for NML network. Preferred and non-preferred objects were identified from electrophysiological responses recorded in the NML network of monkey M1 (Fig. 2a, top) by computing average, baseline-subtracted responses in the window [60 220] ms after stimulus onset (the baseline was computed from the window [−25 25] ms), averaging across all 24 views. The localizer contained three types of block. Block 1 contained images of the five most-preferred objects each at eight views (0° rotation in the y–z space, first row in Extended Data Fig. 2b). Block 2 contained images of the five least-preferred objects each at eight views. Block 3 contained images of five objects that belonged to the animal category each at eight views. A block containing phase-scrambled noise patterns preceded each stimulus block (using the images shown in blocks 1–3). To construct phase-scrambled images, we performed fast Fourier transform (FFT) on images, added a random phase to each frequency component, and then performed an inverse FFT. During the fMRI experiment, stimuli were presented in 24-s blocks at an interstimulus interval of 500 ms. In each scan, the order of the stimulus blocks was fixed as follows: preferred objects, non-preferred objects, animals, non-preferred objects, animals, preferred objects, animals, preferred objects, non-preferred objects. In addition, a block containing phase-scrambled noise was added at the end of each scan. Each scan lasted 456 s. Four monkeys were tested with this localizer, and 6–9 scans were performed for each monkey.

Localizer for body network. The localizer contained eight types of block, each consisting of 16 images taken from the following 8 categories: monkey bodies, animals, faces, fruits, hands, man-made objects, houses, and scenes. Stimuli were presented in 24-s blocks at an interstimulus interval of 500 ms. In each run, the eight blocks were each presented once, interleaved with phase-scrambled noise patterns (computed using images from the eight object blocks). A block containing phase-scrambled noise was added at the end of each scan. Each scan lasted 408 s. Four monkeys were tested with this localizer, and 6–9 scans were performed for each monkey.

Localizer for stubby network. The localizer contained four types of block, each consisting of 20 images taken from the four quadrants of object PC1–PC2 space (Fig. 4a). The images were selected from an image set containing 19,300 background-free object images (http://www.freepngs.com). The images were passed through AlexNet, and projected to object PC1–PC2 space built using the original 1,224 images (see 'Building an object space using a deep network'). Then 20 different images were selected from each of the four quadrants of object PC1–PC2 space, each with a polar angle roughly centred on the respective quadrant. The images were presented in 24-s blocks at an interstimulus interval of 500 ms. In each run, the four blocks were each presented twice, interleaved with phase-scrambled noise patterns (computed using images from the four object blocks). A block containing phase-scrambled noise was added at the end of each scan. Each scan lasted 408 s. Four monkeys were tested with this localizer, and 6–18 scans were performed for each monkey.

Localizer for face network. The localizer contained five types of block, consisting of faces, hands, technological objects, vegetables/fruits, and bodies. Face blocks were presented in alternation with non-face blocks. Stimuli were presented in 24-s blocks at an interstimulus interval of 500 ms. In each run, the face block was repeated four times and each of the non-face blocks was shown once. Blocks of grid-scrambled noise patterns preceded each stimulus block. A block containing grid-scrambled noise was added at the end of each scan. Each scan lasted 408 s. Additional details were as described previously[32]. Four monkeys were tested with this localizer, and 5–12 scans were performed for each monkey.

Localizer for scene network. The localizer contained ten types of block: five scene blocks and five non-scene blocks. Stimuli were presented in 24-s blocks at an interstimulus interval of 500 ms. In each run, the ten blocks were each presented once, interleaved with blocks of grid-scrambled noise. Additional details were as previously described[5]. Two monkeys were tested with this localizer, and 8–12 scans were performed for each monkey.

Localizer for colour network. The localizer contained two types of block: a colour block and a grey block. The colour block consisted of an equiluminant red/green colour grating (2.9 cycles/degree, drifting at 0.75 cycles/s), while the grey block consisted of an identical black–white grating. Stimuli were presented in 24-s blocks, 16 blocks to a run. Each scan lasted 432 s. Additional details were as previously described[8,33]. Four monkeys were tested with this localizer, and 8–14 scans were performed for each monkey.

Localizer for 3D network. The 3D localizer contained two sets of blocks. One set of blocks contained 3D shapes generated by random dot stereograms, including curved shapes such as ripples and saddles and simple flat shapes such as stars and squares. The other set of blocks contained random dots presented at zero disparity. The two sets of blocks were interleaved, and each block lasted 24 s. The images were presented at an interstimulus interval of 500 ms. Each scan lasted 600 s. Monkeys viewed the stimuli through red–green glasses. Four monkeys were tested with this localizer, and 5–12 scans were performed for each monkey.

Silhouette experiment. The localizer contained four types of block, each consisting of 20 images taken from the four quadrants of object PC1–PC2 space (Extended Data Fig. 6c). The images were selected from an image set containing 19,300 background-free object images (images from http://www.freepngs.com). The images were first binarized by setting any pixel that belonged to the object to 0 and any pixel that did not belong to the object to 1. Images were then passed through AlexNet and projected to object PC1–PC2 space built using the original 1,224 images (see ‘Building an object space using a deep network’). Then, 20 different images were selected from each of the four quadrants of object PC1–PC2 space. The images were presented in 24-s blocks at an interstimulus interval of 500 ms. In each run, the four blocks were each presented twice, interleaved with blocks only showing a background with fixation point. A block containing a background with fixation point was added at the end of each scan. Each scan lasted 408 s. Three monkeys were tested with this localizer, and 12–24 scans were performed for each monkey.

Fake object experiment. The experiment was largely identical to the silhouette experiment, but with different stimuli. We used a deep GAN[24] to generate ‘fake object’ images (Extended Data Fig. 6d). The GAN was trained to generate images using response patterns in AlexNet layer fc6. To generate fake objects, we first passed an image set containing 19,300 real object images through Alexnet; for each object image, a 4,096-unit response pattern for layer fc6 was generated. We randomly selected pairs of different patterns, and evenly and randomly recombined these pairs into new patterns[34]. Each new pattern was passed into the GAN to generate one fake object image. Twenty thousand new ‘fake objects’ were generated, and four groups of stimuli (twenty images per group) were selected from this set on the basis of their projection onto PC1–PC2 space. Three monkeys were tested with this localizer, and 10–32 scans were performed for each monkey.

Deep dream experiment. The experiment was largely identical to the silhouette experiment, but with different stimuli. We used deep dream techniques (Matlab 2017b, Deep Learning Toolbox, deepdreamImage function) to generate images projecting strongly onto the four quadrants of object space. Instead of performing gradient ascent on activity of a single fc6 unit, four groups of images were generated through gradient ascent on activation of four fictive units (PC1 + PC2, PC1 − PC2, −PC1 − PC2, −PC1 + PC2), corresponding to linear weighted sums of fc6 units (Extended Data Fig. 6e). For each fictive unit, 20 different images were generated after 100 iterations of gradient ascent, starting with different Gaussian noise patterns. We further confirmed that the images projected to extreme coordinates in PC1–PC2 space by passing the images through AlexNet and projecting the resulting fc6 response pattern onto PC1–PC2 space. Three monkeys were tested with this localizer, and 12–22 scans were performed for each monkey.

fMRI scanning and analysis

Five male rhesus macaques were trained to maintain fixation on a small spot for a juice reward. Eye position was monitored using an infrared camera (ISCAN) sampled at 120 Hz. Monkeys were scanned in a 3T TIM (Siemens, Munich, Germany) magnet equipped with AC88 gradient insert while passively viewing images on a screen. Feraheme contrast agent was injected to improve the signal/noise ratio for functional scans. A single-loop coil was used for structural scans at isotropic 0.5 mm resolution. A custom eight-channel coil was used for functional scans at isotropic 1 mm resolution. Further details about the scanning protocol were as described previously[35].

MRI data analysis. Surface reconstruction based on anatomical volumes was performed using FreeSurfer[36] after skull stripping using FSL’s Brain Extraction Tool (University of Oxford). After applying these tools, segmentation was further refined manually.

Analysis of functional volumes was performed using the FreeSurfer Functional Analysis Stream[37]. Volumes were corrected for motion and undistorted based on acquired field map. The resulting data were analysed using a standard general linear model. For the scene contrast, the average of all scene blocks was compared to the average of all non-scene blocks. For the face contrast, the average of all face blocks was compared to the average of all non-face blocks. For the colour contrast, the colour block was compared to the non-colour blocks. For the body contrast, monkey body and animal blocks were compared to all other blocks. For the stubby contrast, the stubby, inanimate object block was compared to three other blocks. For the 3D contrast, the 3D shape blocks were compared to zero disparity blocks. For the microstimulation contrast, blocks with concomitant electrical stimulation were compared to blocks without stimulation. All the contrasts were performed with a non-paired two-sided *t*-test. *P* value was not adjusted for multiple comparisons.

To determine the area of TE and TEO in each subject, we first co-registered the MRI volume for each subject to a monkey atlas[38]. Then each subject’s TE and TEO were defined using the atlas.

To quantify the reproducibility of patch progression on the cortical surface, we plotted significance values for the four stimulus comparisons defining the four networks in Fig. 4c along three paths in posterior, middle, and anterior IT tracing the centre of the grey matter, spanning the following ranges: 1) lower bank of STS and inferotemporal gyrus at AP position 3; 2) lower bank of STS and inferotemporal gyrus at AP position 13; 3) antero-dorsal (TEad) and antero-ventral (TEav) parts of area TE at AP position 18. Non-significant responses ($P > 10^{-3}$) were set to 0.

Microstimulation

The stimulation protocol followed a block design. We interleaved nine blocks of fixation-only with eight blocks of fixation plus electrical microstimulation; we started and ended with a block without microstimulation. Each block lasted 32 s. During microstimulation blocks we applied one pulse train per second, lasting 200 ms with a pulse frequency of 300 Hz. Bipolar current pulses were charge balanced, with a phase duration of 300 µs and a distance between the two phases of 150 µs. We used a current amplitude of 300 µA. Stimulation pulses were delivered using a computer-triggered pulse generator (S88X; Grass Technologies) connected to a stimulus isolator (A365, World Precision Instruments). All stimulus generation equipment was stored in the scanner control room; the coaxial cable was passed through a wave guide into the scanner room. We obtained 30 scans for monkey M1.

Single-unit recording

Tungsten electrodes (1–20 MΩ at 1 kHz, FHC) were back-loaded into plastic guide tubes. The guide tube length was set to reach approximately 3–5 mm below the dura surface. The electrode was advanced slowly using a manual advancer (Narishige Scientific Instrument, Tokyo, Japan). Neural signals were amplified and extracellular action potentials were isolated using the box method in an on-line spike sorting system (Plexon, Dallas, TX, USA). Spikes were sampled at 40 kHz. All spike data were re-sorted using off-line spike sorting clustering algorithms (Plexon). We recorded data from every neuron encountered. Only well-isolated units were considered for further analysis; otherwise, every neuron was included for analysis. Electrodes were lowered through custom angled grids that allowed us to reach the desired targets; custom software was used to design the grids and plan the electrode trajectories[39].

Behavioural task

Monkeys were head fixed and passively viewed the screen in a dark Wisconsin box. Stimuli for electrophysiology were presented on a CRT monitor (DELL P1130). The screen size covered 27.7 × 36.9 visual degrees and stimulus size spanned 5.7°. The fixation spot size was 0.2° in diameter. Images were presented in random order using custom software. Eye position was monitored using an infrared eye tracking system (ISCAN). Juice reward was delivered every 2–4 s if fixation was properly maintained.

Data analysis

Computing view-identity similarity matrices. For each network, we first identified the 11 most-preferred objects by computing average, baseline-subtracted responses in the window [60 220] ms after stimulus onset (the baseline was computed from the window [−25 25] ms), averaging across all 24 views. We then used responses to these 11 most-preferred objects at 24 views (264 images in total) for the analysis. A 264 × 264 similarity matrix of Pearson's correlation coefficients was computed between the population response vector from each patch to each of the 264 stimuli. Owing to size limitations, only the first 88 × 88 (first 8 views) are shown in Fig. 3a. To compute view-invariant identity selectivity as a function of time (Extended Data Fig. 3f), at each time point t between 0 and 400 ms following stimulus onset, in increments of 50 ms, a similarity matrix was computed from mean responses between $t - 25$ and $t + 25$ ms. We then calculated a 'same object correlation value' as the average of correlation values between the same object across different views (solid traces in Extended Data Fig. 3f), and a 'different object correlation value' as the average of correlation values between different objects across same and different views (dashed traces in Extended Data Fig. 3f).

Building an object space using a deep network. The stimulus set consisting of 51 objects at 24 different views (1,224 images) was fed into the pre-trained network AlexNet[6]. Then the responses of 4,096 nodes in layer fc6 were extracted to form a 1,224 × 4,096 matrix. PCA was performed on this matrix, yielding 1,223 PCs, each of length 4,096. To further reduce the dimensionality of the object space, we retained only the first 50 PCs, which captured 85% of the response variance across AlexNet fc6 units. The first two dimensions accounted for 27% of the response variance across AlexNet fc6 units.

To test the robustness of object PC1–PC2 space to the particular set of 1,224 images used to build it (Extended Data Fig. 4d, e), over multiple iterations we randomly picked 1,224 images from a new database (http://www.freepng.com) containing 19,300 background-free object images. The 1,224 images were fed into Alexnet, and we followed the same procedure to build a new object space, which we call PC1′–PC2′ space. The original 1,224 images were passed through Alexnet, and the vector of fc6 unit activations was projected onto both PC1–PC2 space and PC1′–PC2′ space. Thus we have a set of 1,224 coordinates in both PC1–PC2 space and PC1′–PC2′ space. We then determined the best affine transform of PC1′–PC2′ space so that the coordinates of the 1,224 images in the two spaces would have minimum distance using linear regression.

$$\begin{pmatrix} x_{1,1} & x_{1,2} \\ & \vdots \\ x_{1224,1} & x_{1224,2} \end{pmatrix} = \begin{pmatrix} x'_{1,1} & x'_{1,2} \\ & \vdots \\ x'_{1224,1} & x'_{1224,2} \end{pmatrix} \begin{pmatrix} u & w \\ u' & w' \end{pmatrix}$$

where $(x_{i,1}\, x_{i,2})$ is the coordinate of image i in PC1–PC2 space, and $(x'_{i,1}\, x'_{i,2})$ is the coordinate of image i in PC1′–PC2′ space. After matching, we calculated the Pearson's correlation r between PC1 and affined transformed PC1′, and PC2 and affine transformed PC2′. We used a similar procedure to test the robustness of object PC1–PC2 space to the particular network used to compute it (Extended Data Fig. 4i).

Quantifying the aspect ratio of objects. The aspect ratio of an object (Extended Data Fig. 3g) was defined as a function of perimeter P and area A:

$$\text{Aspect ratio} = \frac{P^2}{4\pi A}$$

P was measured by the number of pixels lying on the object image's boundary, and was computed using the Matlab *bwboundaries* function. The area was measured by the number of pixels that belonged to the object, and was computed using the Matlab *regionsprops* function.

Computing the preferred axis of an IT cell. The number of spikes in a time window of 60–220 ms after stimulus onset was counted for each stimulus. To estimate the preferred axis, we used linear regression to compute the coefficients **c** in the equation $\mathbf{R} = \mathbf{c} \cdot \mathbf{F} + c_0$, where **R** is the response vector of the cell to the set of images, **F** is the matrix of 50D object feature vectors for the set of images, and c_0 is a constant offset. Using this definition of preferred axis, cells will necessarily show an increasing firing rate for increasing value of projection onto the preferred axis. To generate Fig. 3c, we randomly picked half the stimulus trials to compute the preferred axis for each cell, and then used the held-out data to plot the responses shown.

Computing tuning along dimensions orthogonal to the preferred axis. To compute tuning along orthogonal dimensions (Extended Data Fig. 5e, black traces), for each neuron we first computed the preferred axis. There are 49 dimensions spanning the subspace orthogonal to this preferred axis. To find the longest orthogonal axis in this 49D subspace, we first represented each of the 1,224 images in our stimulus set as a 50D vector in object space, and subtracted the preferred axis of the cell from each of these image feature vectors, to obtain a set of feature vectors lying in the 49D orthogonal subspace. We performed PCA on this set of 1,224 vectors, and picked the top PC. This PC represents the axis orthogonal to the preferred axis of the cell that captures the largest variation in the images. For each cell, the tuning curve of the cell along this axis was computed.

Quantifying consistency of a cell's preferred axis. The consistency of the preferred axis of each cell (Extended Data Fig. 5a) was measured as follows: in each iteration, the whole image set (1,224 images) was randomly split into two subsets of 612 images, and a preferred axis was calculated using the responses to each subset. Then the Pearson correlation (r) was calculated between the two. This was repeated 100 times, and the consistency of preferred axis for the cell was defined as the average r value across 100 iterations.

Quantifying explained variance along an object dimension. In Extended Data Fig. 11b, c, the explained variance R^2 was determined by the difference between the reconstructed feature value y'_i and the real object feature value y_i:

$$R^2 = 1 - \frac{\sum_1^{1224}\left(y_i - y'_i\right)^2}{\sum_1^{1224}\left(y_i - \bar{y}\right)^2}$$

Quantifying explained variance in single neuron firing rate and model comparison. In Extended Data Fig. 5b–d, to compute explained variance we first fit responses to a set of 1,593 objects (Extended Data Fig. 2d) using the axis model and then tested it on responses to a different set of 100 objects. To obtain high signal quality, the 100 objects were repeated 15–30 times. In Extended Data Fig. 5c, d, we compared three different models: (1) the axis model, which assumed the 50D features are combined linearly; (2) a Gaussian model, defined as $R = ae^{\left(-(\mathbf{x}-\mathbf{x}_0)^2/\sigma^2\right)}$; and (3) a quadratic model, defined as $R = a(\mathbf{x} - \mathbf{x}_0)^2 + b(\mathbf{x} - \mathbf{x}_0) + c$. The percentage of explainable variance in responses to 100 objects explained by each model was used to quantify the quality of fit. In Extended Data Fig. 5b, for each cell the explained variance R^2 was determined by the difference between the predicted responses r'_i and real observed responses to 100 test images r_i:

$$R^2 = 1 - \frac{\sum_1^{100}(r_i - r'_i)^2}{\sum_1^{100}(r_i - \bar{r})^2}$$

For calculating the upper bound of explained variance (*y*-axis values in Extended Data Fig. 5b), different trials of responses to the stimuli were randomly split into two halves. The Pearson correlation (r) between the average responses from two half-splits across images was calculated and corrected using the Spearman–Brown correction:

$$r' = \frac{2r}{1+r}$$

The square of r' was considered as the upper bound of the explained variance.

k-means cluster analysis. To determine whether neurons in the same network are grouped as a cluster based on their preferred axes, we applied k-means analysis on the entire population of neurons recorded in the four networks (Extended Data Fig. 8g, h). The distance between each pair of neurons was calculated as the Pearson's correlation between preferred axes of the neurons in the 50D space. To determine the optimal number of clusters, we calculated the Calinski–Harabasz value (CH) for different numbers of clusters (k).

$$\mathrm{CH}(k) = \frac{B(k)\times(n-k)}{w(k)\times(k-1)}$$

$B(k)$ is the between-cluster variation, $w(k)$ is the within-cluster variation, n is the number of neurons, and k is the cluster number. The larger the value of CH, the better the cluster model is. To check whether clusters exist beyond the first two PCs, k-means analysis was performed by defining the distance between a pair of neurons as the correlation in preferred axes in 48 dimensions after removing the first two PCs in the original 50D object space.

Decoding analysis. We found that cells in each IT network were performing linear projection onto specific preferred axes (Fig. 3c, Extended Data Fig. 5a, e) and could be well modelled by the equation $\mathbf{R} = \mathbf{c}\cdot\mathbf{f} + \mathbf{c}_0$, where $\mathbf{R}$ is the vector of responses of different neurons, $\mathbf{c}$ is the matrix of weighting coefficients for different neurons, $\mathbf{f}$ is the vector of feature values in the object space, and $\mathbf{c}_0$ is the offset vector. This suggests that by simply inverting this equation, we should be able to decode the vector of feature values in the object space from the IT response vector: $\mathbf{f} = \mathbf{R}\cdot\mathbf{c}' + \mathbf{c}_0'$. We first used responses to all but one of the objects (1,224 − 24 = 1,200 images) to fit $\mathbf{c}'$ and $\mathbf{c}_0'$. Then the linear model was applied to responses to the remaining object for each of the 24 views to compute the predicted feature vector (Fig. 5, Extended Data Fig. 11).

To quantify overall decoding accuracy (Extended Data Fig. 11d–f), we randomly selected a subset of N object images from the set of 1,224 images and compared their actual object feature vectors to the reconstructed feature vector for one image ('target') in the set of 1,224 using Euclidean distance. If the object feature vector with the smallest distance to the reconstructed object feature vector portrays the actual target, the decoding is considered correct. We repeated the procedure 100 times for each of the 1,224 object images to estimate decoding accuracy.

Object reconstruction. To reconstruct objects from neural activity (Fig. 5), we used a pre-trained GAN[24]. For each image, a 50D object feature vector was reconstructed from neural activity elicited by that image; then the resulting 50D feature vector was transformed back into an fc6 layer pattern using the Moore–Penrose pseudoinverse. Finally, we passed this fc6 response pattern to the generative network to generate reconstructed images. Since the generative network cannot perfectly reconstruct images from AlexNet fc6 layer responses, for comparison we also reconstructed each image using (1) its original fc6 response pattern and (2) the original fc6 response pattern projected onto the 50D object space; the latter constitutes the best possible reconstruction. We computed a 'normalized distance' to quantify the reconstruction accuracy for each object:

$$\text{Normalized distance} = \frac{\left|\mathbf{fc6}_{\text{recon}} - \mathbf{fc6}_{\text{original}}\right|}{\left|\mathbf{fc6}_{\text{best possible recon}} - \mathbf{fc6}_{\text{original}}\right|},$$

Where $\mathbf{fc6}_{\text{recon}}$ is the fc6 response pattern to the reconstruction obtained using neural data, $\mathbf{fc6}_{\text{original}}$ is the fc6 response pattern to the original image shown to the monkey and $\mathbf{fc6}_{\text{best possible recon}}$ is the fc6 response pattern to the best possible reconstruction.

As an alternative to directly reconstructing images using a GAN, we recovered images using an auxiliary database (Extended Data Fig. 11g, h). We passed an image set containing 18,700 background-free object images (http://www.freepngs.com) and 600 face images (FEI database), none of which had been shown to the monkey, through AlexNet, and projected these images to the object space computed using our original stimulus set of 1,224 images. For each image, the object feature vector reconstructed from neural activity was compared with object feature vectors for images from the new image set. The image in the new image set with the smallest Euclidean distance to the reconstructed object feature vector was considered as the 'reconstruction' of this object feature vector.

To take into account the fact that the object images used for reconstruction did not include any of the object images shown to the monkey, setting a limit on how good the reconstruction can be, we computed a 'normalized distance' to quantify the reconstruction accuracy for each object. We defined the normalized reconstruction distance for an image as

$$\text{Normalized distance} = \frac{\left|\mathbf{v}_{\text{recon}} - \mathbf{v}_{\text{original}}\right|}{\left|\mathbf{v}_{\text{best possible recon}} - \mathbf{v}_{\text{original}}\right|},$$

where $\mathbf{v}_{\text{recon}}$ is the feature vector reconstructed from neuronal responses, $\mathbf{v}_{\text{original}}$ is the feature vector of the image presented to the monkey, and $\mathbf{v}_{\text{best possible recon}}$ is the feature vector of the best possible reconstruction. A normalized distance of one means that the reconstruction has found the best solution possible.

Object specialization index computation. To quantify whether a particular object is better represented by a particular network compared to other networks (Extended Data Fig. 11i), for each of 1,224 objects and each of three networks (body, NML, stubby), we computed a specialization index SI_{ij} that measures how much better decoding accuracy for object i computed from activity in network j is compared to decoding accuracy for object i computed across all other networks using the same number of neurons:

$$\mathrm{SI}_{ij} = \frac{\mathrm{DA}_{i,j} - \mathrm{DA}_{i,\sim j}}{\mathrm{DA}_{i,j} + \mathrm{DA}_{i,\sim j}},$$

where $\mathrm{DA}_{i,j}$ is the decoding accuracy for object i computed using N random neurons from network j, and $\mathrm{DA}_{i,\sim j}$ is the decoding accuracy for object i computed using N random neurons from all networks except j. SI_{ij} quantifies how specialized network j is for representing object i.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The data that support the findings of this study are available from the lead corresponding author (D.Y.T.) upon reasonable request.

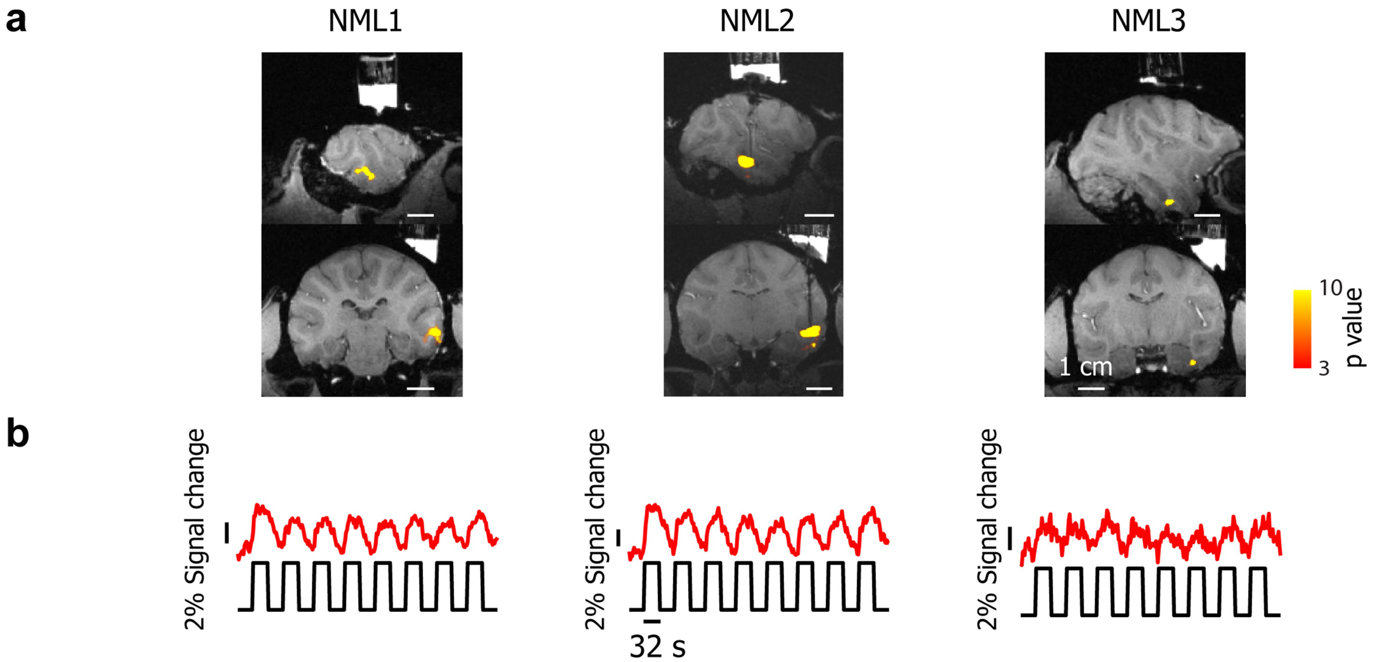
32. Tsao, D. Y., Freiwald, W. A., Tootell, R. B. & Livingstone, M. S. A cortical region consisting entirely of face-selective cells. *Science* **311**, 670–674 (2006).
33. Chang, L., Bao, P. & Tsao, D. Y. The representation of colored objects in macaque color patches. *Nat. Commun*. **8**, 2064 (2017).
34. Ponce, C. R. et al. Evolving images for visual neurons using a deep generative network reveals coding principles and neuronal preferences. *Cell* **177**, 999–1009. e1010 (2019).
35. Ohayon, S., Grimaldi, P., Schweers, N. & Tsao, D. Y. Saccade modulation by optical and electrical stimulation in the macaque frontal eye field. *J. Neurosci*. **33**, 16684–16697 (2013).
36. Dale, A. M., Fischl, B. & Sereno, M. I. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* **9**, 179–194 (1999).
37. Reuter, M. & Fischl, B. Avoiding asymmetry-induced bias in longitudinal image processing. *Neuroimage* **57**, 19–21 (2011).
38. Reveley, C. et al. Three-dimensional digital template atlas of the macaque brain. *Cereb. Cortex* **27**, 4463–4477 (2017).
39. Ohayon, S. & Tsao, D. Y. MR-guided stereotactic navigation. *J. Neurosci. Methods* **204**, 389–397 (2012).
40. Konkle, T. & Caramazza, A. Tripartite organization of the ventral stream by animacy and object size. *J. Neurosci*. **33**, 10235–10242 (2013).
41. Long, B., Yu, C.-P. & Konkle, T. Mid-level visual features underlie the high-level categorical organization of the ventral stream. *Proc. Natl Acad. Sci. USA* **115**, E9015–E9024 (2018).

Acknowledgements This work was supported by NIH (DP1-NS083063, R01-EY030650), the Howard Hughes Medical Institute, and the Tianqiao and Chrissy Chen Institute for Neuroscience at Caltech. We thank A. Flores for technical support, and members of the Tsao laboratory, N. Kanwisher, A. Kennedy, S. Kornblith, and A. Tsao for critical comments.

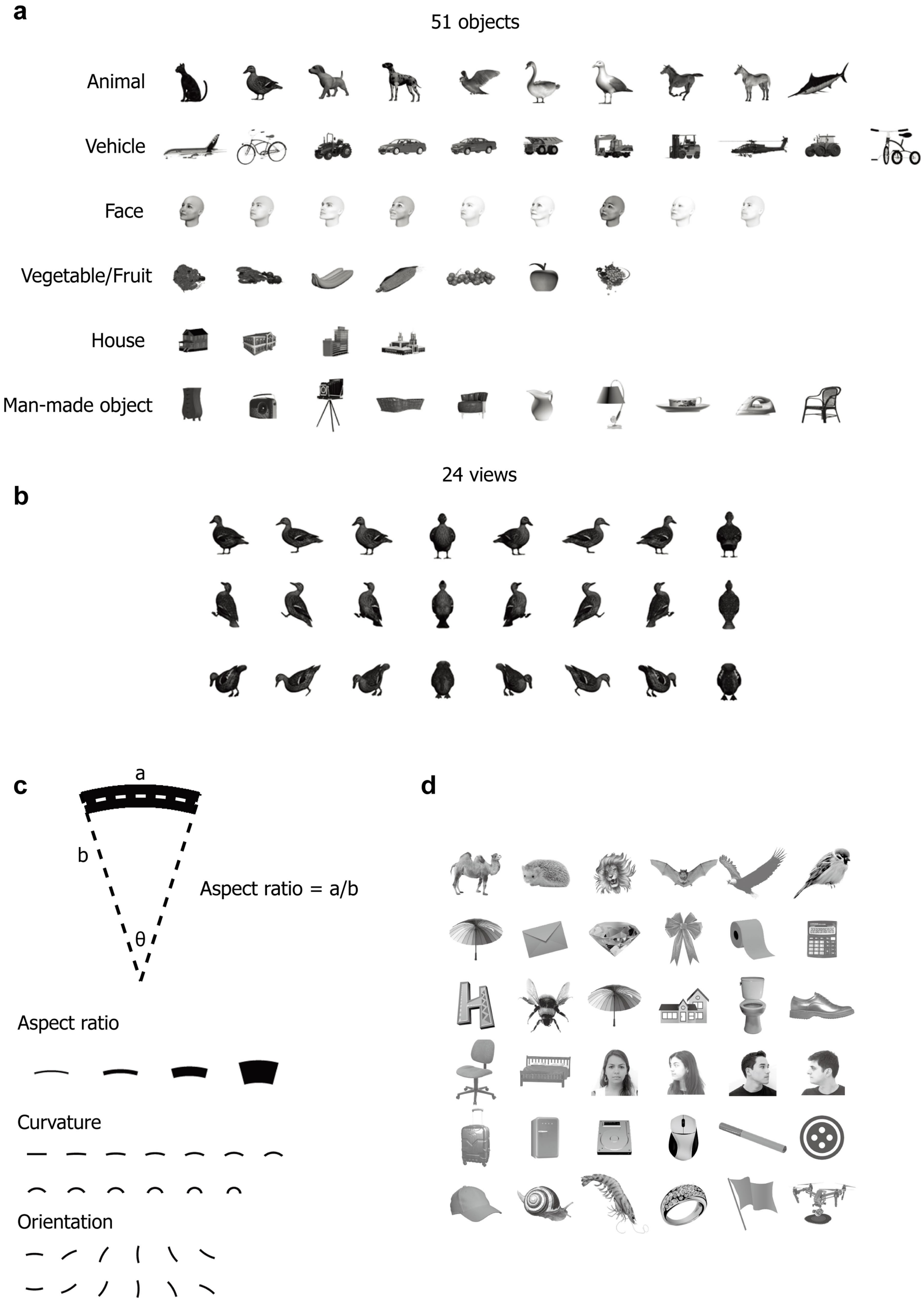
Author contributions P.B. and D.Y.T. designed the experiments, P.B. and L.S. collected the data, and P.B. analysed the data. M.M. provided technical advice on neural networks. P.B. and D.Y.T. interpreted the data and wrote the paper.

Competing interests The authors declare no competing interests.

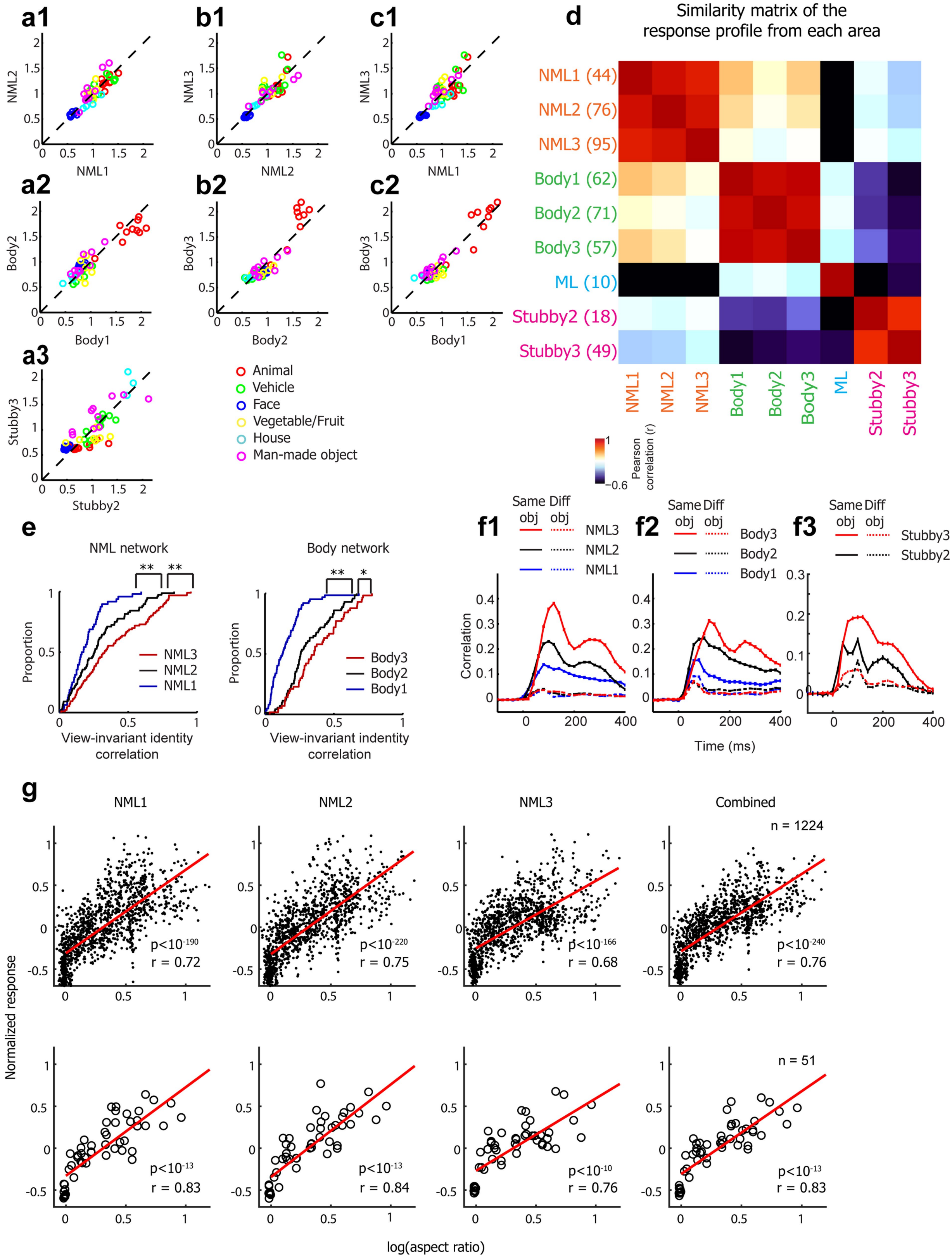
Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2350-5.
Correspondence and requests for materials should be addressed to P.B. or D.Y.T.
Reprints and permissions information is available at http://www.nature.com/reprints.



Extended Data Fig. 1 | Time courses from NML1-3 during microstimulation of NML2. **a**, Sagittal (top) and coronal (bottom) slices showing activation in response to microstimulation of NML2. Dark track shows electrode targeting NML2. **b**, Time courses of microstimulation (black) together with fMRI response (red) from each of the three patches of the NML network.

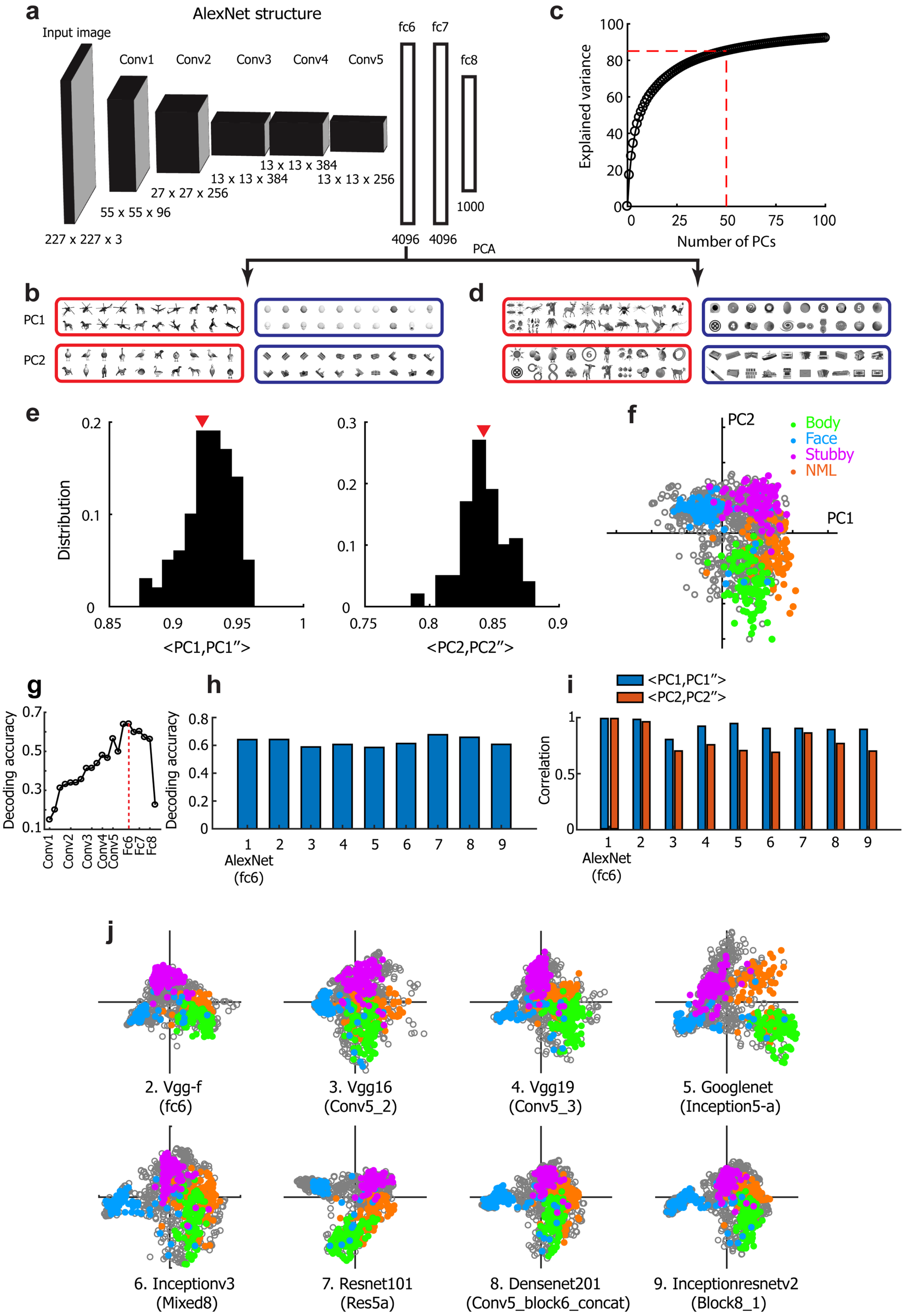


Extended Data Fig. 2 | Stimuli used in electrophysiological recordings. **a**, Fifty-one objects from six categories were shown to monkeys. **b**, Twenty-four views for one example object, resulting from rotations in the x–z plane (abscissa) combined with rotations in the y–z plane (ordinate). **c**, A line segment that was parametrically varied along three dimensions was used to test the hypothesis that cells in the NML network are selective for aspect ratio: 4 aspect ratio levels × 13 curvature levels × 12 orientation levels. **d**, Thirty-six example object images from an image set containing 1,593 images.



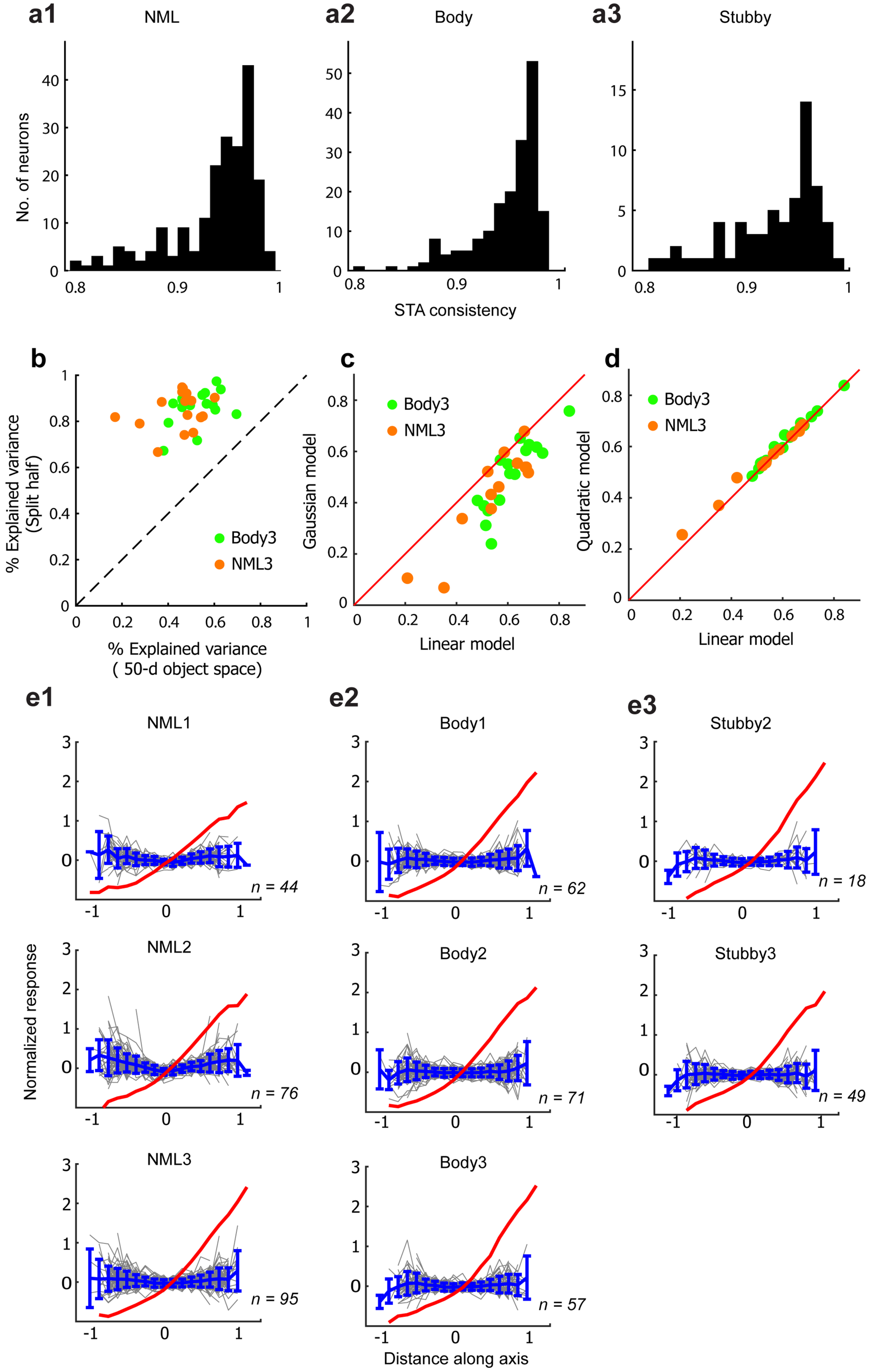
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Additional neuronal response properties from different patches. **a1**, Average responses to 51 objects across all cells from patch NML2 plotted against those from patch NML1. The response to each object was defined as the average response across 24 views and across all cells recorded from a given patch. **b1**, As in **a1** for NML3 against NML2. **c1**, As in **a1** for NML3 against NML1. **a2**, **b2**, **c2**, As in **a1**, **b1**, **c1** for three patches of the body network. **a3**, As in **a1** for Stubby3 against Stubby2. **d**, Similarity matrix showing the Pearson correlation values (*r*) between the average responses to 51 objects from 9 patches across 4 networks. **e**, Left, cumulative distributions of view-invariant identity correlations for cells in the three patches of the NML network. Right, as on left for cells in the three patches of the body network. For each cell, the view-invariant identity correlation was computed as the average across all pairs of views of the correlation between response vectors to the 51 objects at a pair of distinct views. The distribution of view-invariant identity correlations was significantly different between NML1 and NML2 (two-tailed *t*-test, $P < 0.005$, $t(118) = 2.96$), NML2 and NML3 (two-tailed *t*-test, $P < 0.005$, $t(169) = 2.9$), Body1 and Body2 (two-tailed *t*-test, $P < 0.0001$, $t(131) = 6.4$), and Body2 and Body3 (two-tailed *t*-test, $P < 0.05$, $t(126) = 2.04$). *$P < 0.05$, **$P < 0.01$. **f1**, Time course of view-invariant object identity selectivity for the three patches in the NML network, computed using responses to 11 objects at 24 views and a 50-ms sliding response window (solid lines). As a control, time courses of correlations between responses to different objects across different views were also computed (dashed lines) (see Methods). **f2**, As in **f1** for body network. **f3**, As in **f1** for stubby network. **g**, Top, average responses to each image across all cells recorded from each patch plotted against the logarithm of aspect ratio of the object in each image (see Methods). Pearson *r* values are indicated in each plot (all $P < 10^{-10}$). The rightmost column shows results with cells from all three patches grouped together. Bottom, As on top, with responses to each object averaged across 24 views, and associated aspect ratios also averaged. The rightmost column shows results with cells from all three patches grouped together.



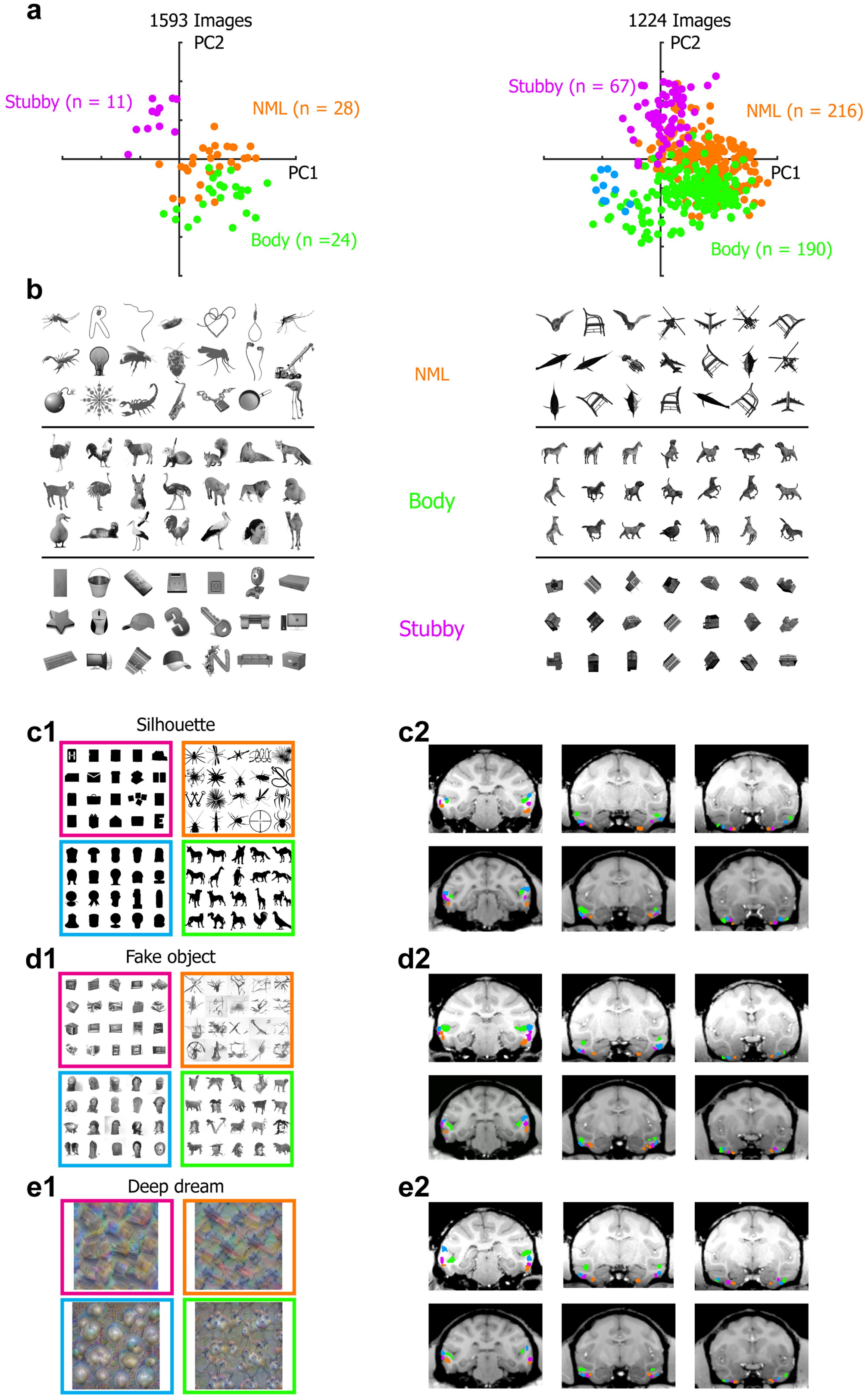
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Building an object space using a deep network. **a**, A diagram illustrating the structure of AlexNet[6]. Five convolution layers are followed by three fully connected layers. The number of units in each layer is indicated below each layer. **b**, Images with extreme values (highest: red, lowest: blue) of PC1 and PC2. **c**, The cumulative explained variance of responses of units in fc6 by 100 PCs; 50 dimensions explain 85% of variance. **d**, Images in the 1,593-image set with extreme values (highest: red, lowest: blue) of PC1 and PC2 built using the 1,593 image set after affine transform (see Methods). Preferred features are generally consistent with those computed using the original image set shown in **b**. However, PC2 no longer clearly corresponds to an animate–inanimate axis; instead, it corresponds to curved versus rectilinear shapes. **e**, Distributions showing the canonical correlation value between the first two PCs obtained by the 1,224-image set and the first two PCs built by other sets of images (1,224 randomly selected non-background object images, left: PC1, right: PC2; see Methods for details). The red triangles indicate the arithmetic mean of the distributions. **f**, We passed 19,300 object images through AlexNet and built PC1–PC2 space using PCA. Then we projected 1,224 images onto this PC1–PC2 space. The top 100 images for each network are indicated by coloured dots (compare Fig. 4b). **g**, Decoding accuracy for 40 images using object spaces built by responses of different layers of AlexNet (computed as in Extended Data Fig. 11d). There are multiple points for each layer because we performed PCA before and after pooling, activation, and normalization functions. Layer fc6 showed the highest decoding accuracy, motivating our use of the object space generated by this layer throughout the paper. **h**, To compare IT clustering determined by AlexNet with that by other deep network architectures, we first identified the layer of each network that gave the best decoding accuracy, as in **g**. The bar plot shows decoding accuracy for 40 images in the 9 different networks using the best-performing layer for each network. **i**, Canonical correlation values between the first two PCs obtained by Alexnet and first two PCs built by eight other deep-learning networks (labelled 2–9). The layer of each network that yielded the highest decoding accuracy for 40 images was used for this analysis. The name of each network and layer can be found in **j**. **j**, As in Fig. 4b using PC1 and PC2 computed from eight other networks.



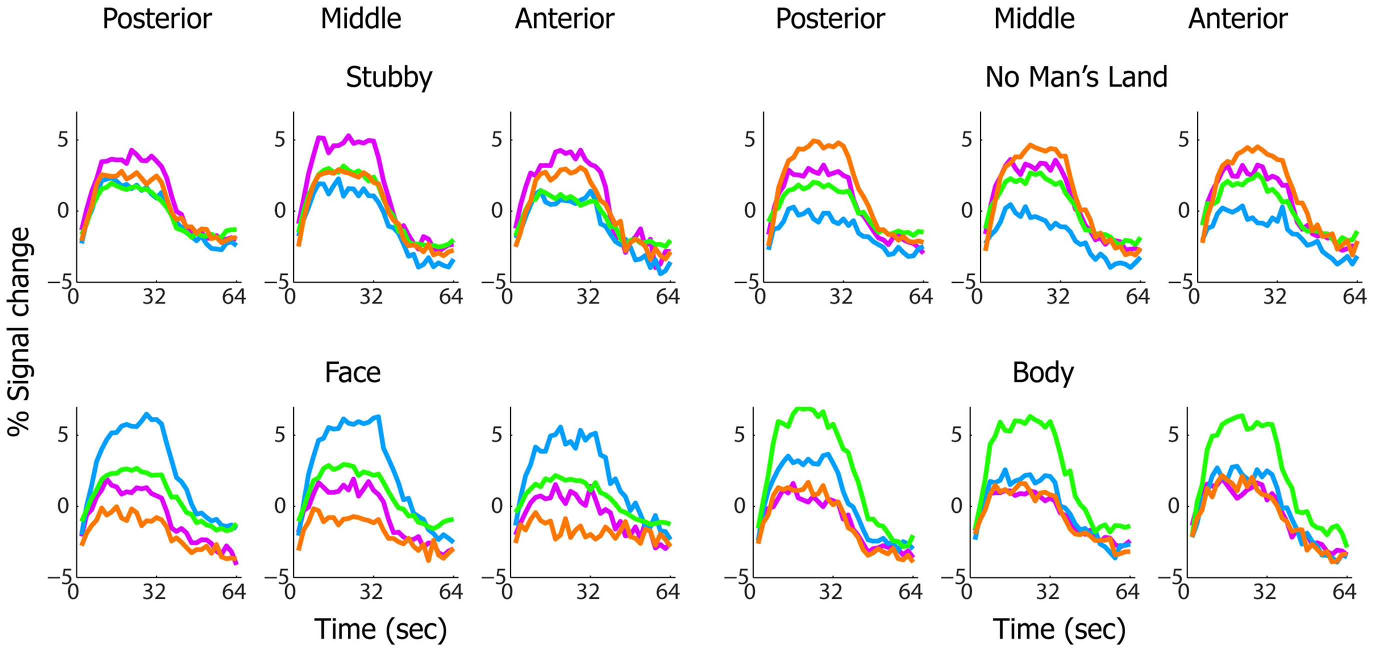
Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Neurons across IT perform axis coding. a1, The distribution of consistency of preferred axis for cells in the NML network (see Methods). **a2**, As in **a1** for the body network. **a3**, As in **a1** for the stubby network. **b**, Different trials of responses to the stimuli were randomly split into two halves, and the average response across half of the trials was used to predict that of the other half. Percentage variances explained, after Spearman–Brown correction (mean 87.8%), are plotted against that of the axis model (mean 49.1%). Mean explainable variance for 29 cells was 55.9%. **c**, Percentage variances explained by a Gaussian model plotted against that of the axis model. **d**, Percentage variances explained by a quadratic model plotted against that of the axis model. Inspection of coefficients of the quadratic model revealed a negligible quadratic term (mean ratio of 2nd-order coefficients/1st-order coefficient, 0.028). **e1**, Top, red line shows the average modulation along the preferred axis across the population of NML1 cells. The grey lines show, for each cell in NML1, the modulation along the single axis orthogonal to the preferred axis in the 50D object space that accounts for the most variability. The blue line and error bars represent the mean and s.d. of the grey lines. Middle, bottom, analogous plots for NML2 and NML3, respectively. **e2**, As in **e1** for the three body patches. **e3**, As in e1 for the two stubby patches.

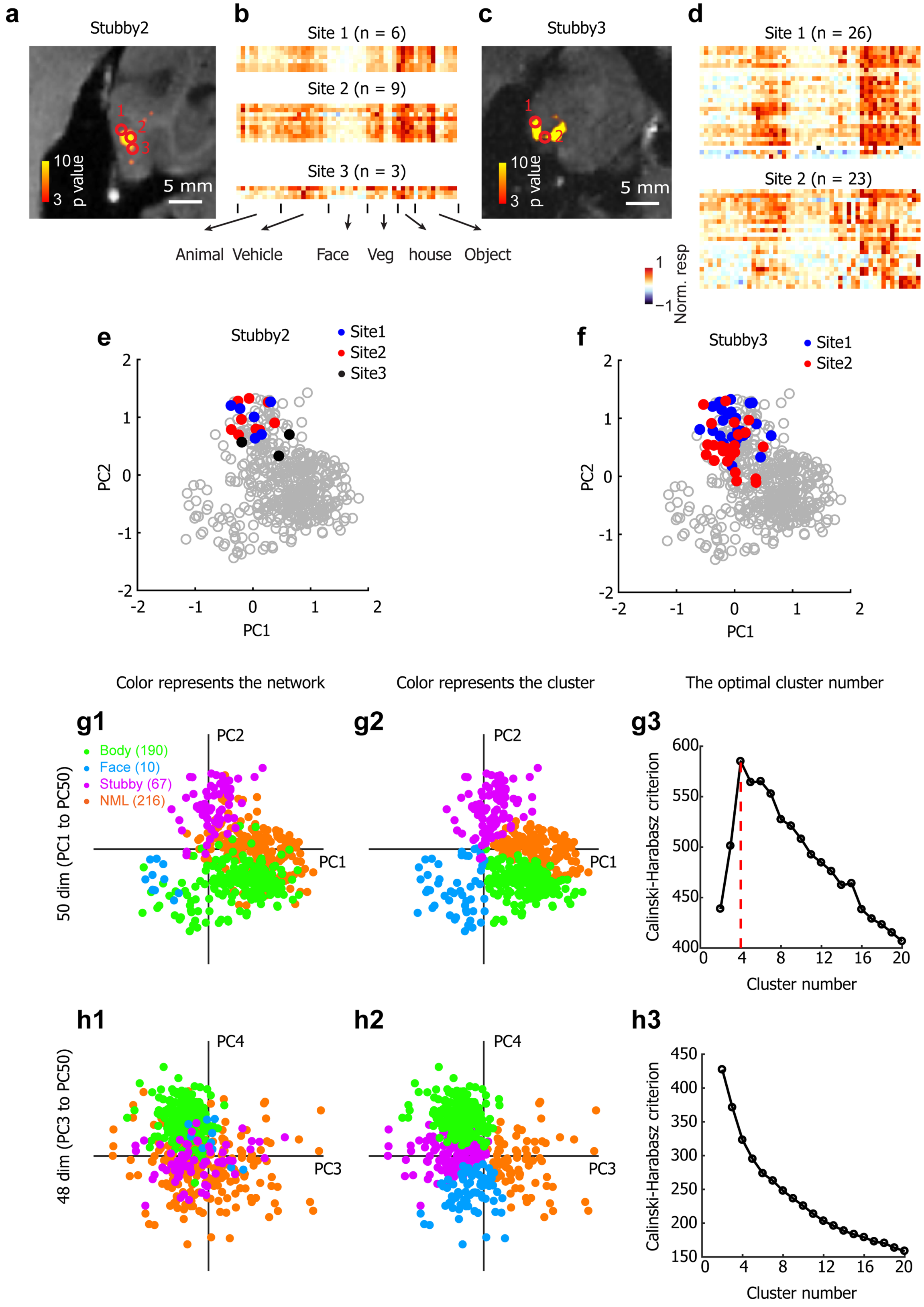


Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Similar functional organization is observed using a different stimulus set. a, Projection of preferred axes onto PC1 versus PC2 for all neurons recorded using two different stimulus sets (left, 1,593 images from freepngs image set; right, the original 1,224 images consisting of 51 objects × 24 views). The PC1–PC2 space for both plots was computed using the 1,224 images. Different colours encode neurons from different networks. **b**, Top 21 preferred stimuli based on average responses from the neurons recorded in three networks to the two different image sets. **c1**, Four classes of silhouette images that project strongly onto the four quadrants of object space. **c2**, Coronal slices from posterior, middle, and anterior IT of monkeys M2 and M3 showing the spatial arrangement of the four networks revealed using the silhouette images in **c1** in an experiment analogous to that in Fig. 4a. **d1**, Four classes of 'fake object' images that project strongly onto the four quadrants of object space. Note that fake objects that project onto the face quadrant no longer resemble real faces. **d2**, As in **c2** with fake object images from **d1**. **e1**, Four example stimuli generated by deep dream techniques that project strongly onto the four quadrants of object space. **e2**, As in **c2** with deep dream images from **e1**. The results in **c**–**e** support the idea that IT is organized according to the first two axes of object space rather than low-level features, semantic meaning, or image organization.

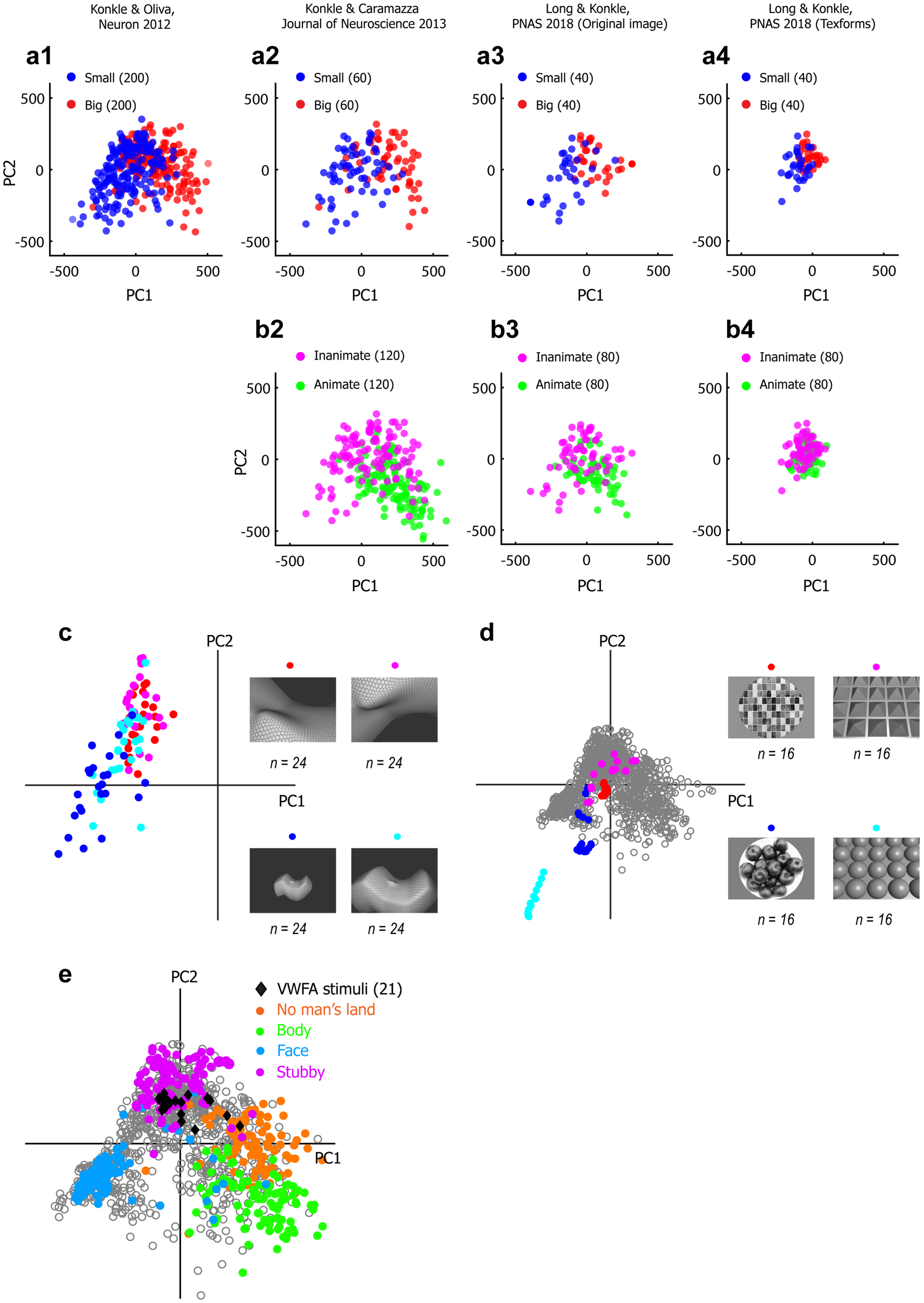


Extended Data Fig. 7 | Response time courses from the four IT networks spanning object space. Time courses were averaged across two monkeys. To avoid selection bias, odd runs were used to identity regions of interest, and even runs were used to compute average time courses from these regions.



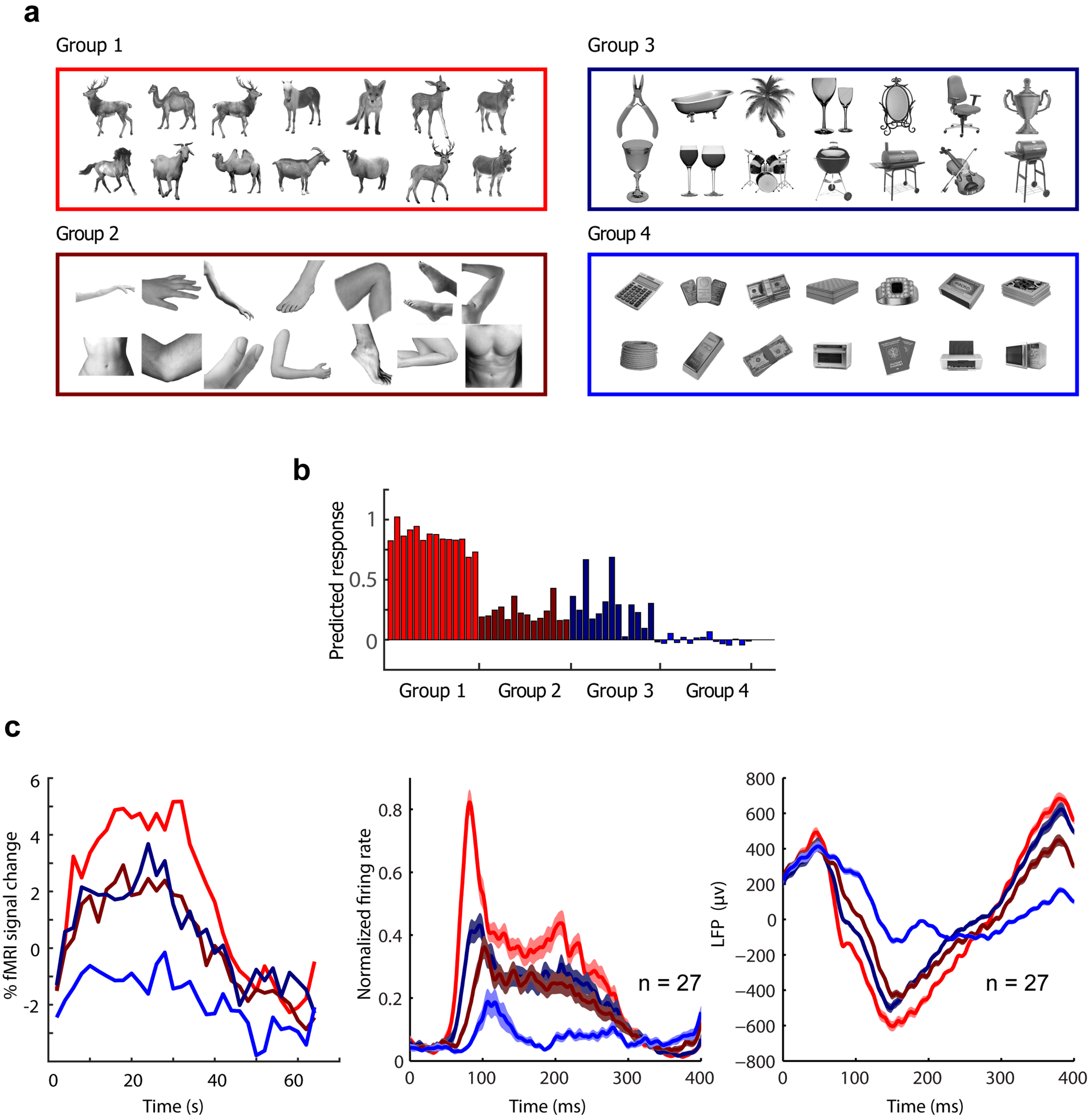
Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Searching for substructure within patches. **a**, Axial view of the Stubby2 patch, together with projections of three recording sites. **b**, Mean responses to 51 objects from neurons grouped by recording sites shown in **a** (same format as Fig. 2a (top)). **c**, Axial view of the Stubby3 patch, together with projections of two recording sites. **d**, Mean responses to 51 objects from neurons grouped by recording sites shown in **c**. **e**, Projection of preferred axis onto PC1–PC2 space for neurons recorded from different sites within the Stubby2 patch. There is no clear separation between neurons from the three sites in PC1–PC2 space. The grey dots represent all other neurons across the four networks. **f**, As in **e** for cells recorded from two sites in the Stubby3 patch. **g1**, Projection of preferred axes onto PC1–PC2 space for all recorded neurons. Different colours encode neurons from different networks. **g2**, As in **g1**, but the colour represents the cluster to which the neurons belong. Clusters were determined by *k*-means analysis, with the number of clusters set to four, and the distance between neurons defined by the correlation between preferred axes in the 50D object space (see Methods). Comparison of **g1** and **g2** reveals highly similarity between the anatomical clustering of IT networks and the functional clustering determined by *k*-means analysis. **g3**, Calinski–Harabasz criterion values were plotted against the number of clusters for *k*-means analysis performed with different numbers of clusters (see Methods). The optimal cluster number is four. **h1**, As in **g1** for projection of preferred axes onto PC3 versus PC4. **h2**, As in **h1**, but the colour represents the cluster to which the neurons belong. Clusters were determined by *k*-means analysis, with the number of clusters set to four, and the distance between neurons defined by the correlation between preferred axes in the 48D object space obtained by removing the first two dimensions. The difference between **h1** and **h2** suggests that there is no anatomical clustering for dimensions beyond the first two PCs. **h3**, As in **g3**, with *k*-means analysis in the 48D object space. By the Calinski–Harabasz criterion, there is no functional clustering for higher dimensions beyond the first two.

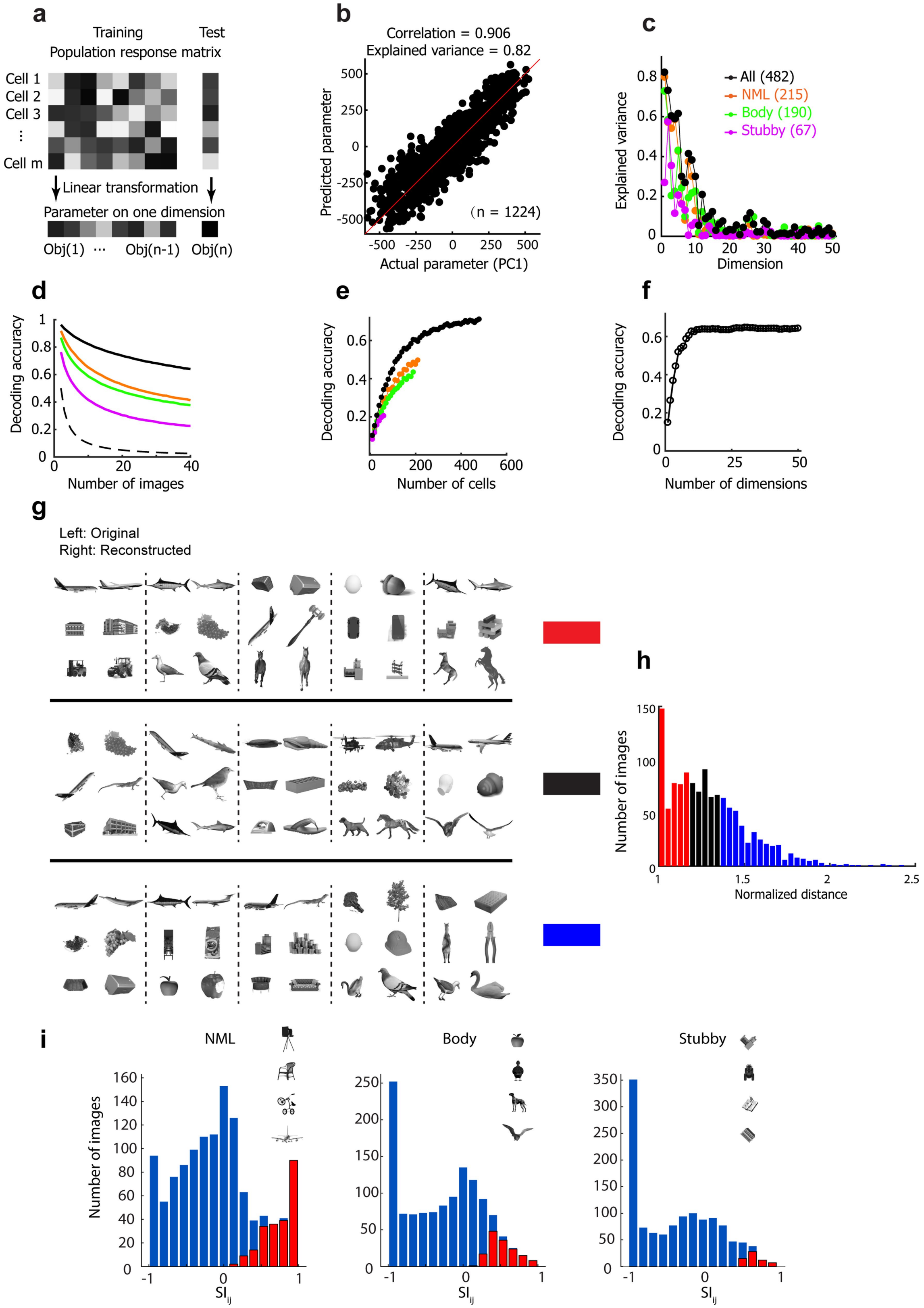


Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | The object space model parsimoniously explains previous accounts of IT organization. a1, The object images used in ref. [14] are projected onto PC1–PC2 space (computed as in Fig. 4b, by first passing each image through AlexNet). A clear gradient from big (red) to small (blue) objects is seen. **a2**, As in **a1**, for the inanimate objects (big and small) used in ref. [40]. **a3**, As in **a1**, for the original object images used in ref. [41]. **a4**, As in **a1**, for the texform images used in ref. [41]. **b2**–**4**, Projection of animate and inanimate images from original object images (**b2**, **b3**) and texforms (**b4**). **c**, Left, coloured dots depict projection of stimuli from the four conditions used in ref. [21]. Right, example stimuli (blue, small object-like; cyan, large object-like; red, landscape-like; magenta, cave-like). **d**, Left, grey dots depict 1,224 stimuli projected onto object PC1–PC2 space; coloured dots depict projection of stimuli from the four blocks of the curvature localizer used in ref. [11]. Right, example stimuli from the four blocks of the curvature localizer (blue, real-world round shapes; cyan, computer-generated 3D sphere arrays; red, real-world rectilinear shapes; magenta, computer-generated 3D pyramid arrays). **e**, Images of English and Chinese words are projected onto object PC1–PC2 space (black diamonds), superimposed on the plot from Fig. 4b. They are grouped into a small region, consistent with their modular representation by the visual word form area.



Extended Data Fig. 10 | Object space dimensions are a better descriptor of response selectivity in the body patch than category labels. **a**, Four classes of stimuli: 1, body stimuli that project strongly onto the body quadrant of object space (bright red); 2, body stimuli that project weakly onto the body quadrant of object space (dark red); 3, non-body stimuli that project equally strongly as group 2 onto the body quadrant of object space (dark blue); and 4, non-body stimuli that project negatively onto the body quadrant of object space (bright blue). **b**, Predicted response of the body patch to each image from the four stimulus conditions in **a**, computed by projecting the object space representation of each image onto the preferred axis of the body patch (determined from the average response of body patch neurons to the 1,224 stimuli). **c**, Left, fMRI response time course from the body patches to the four stimulus conditions in **a**. Centre, mean normalized single-unit responses from neurons in Body1 patch to the four stimulus conditions. Right, mean local field potential from Body1 patch to the four stimulus conditions. Shading represents s.e.



Extended Data Fig. 11 | See next page for caption.

Extended Data Fig. 11 | Object decoding and recovery of images by searching a large auxiliary object database. a, Schematic illustrating the decoding model. To construct and test the model, we used responses of *m* recorded cells to *n* images. Population responses to images from all but one object were used to determine the transformation from responses to feature values by linear regression, and then the feature values of the remaining object were predicted (for each of 24 views). **b**, Model predictions plotted against actual feature values for the first PC of object space. **c**, Percentage explained variances for all 50 dimensions using linear regression based on the responses of four neural populations: 215 NML cells (yellow); 190 body cells (green); 67 stubby cells (magenta); 482 combined cells (black). **d**, Decoding accuracy as a function of the number of object images randomly drawn from the stimulus set for the same four neural populations as in **c**. Dashed line indicates chance performance. **e**, Decoding accuracy for 40 images plotted against different numbers of cells randomly drawn from same four populations as in **c**. **f**, Decoding accuracy for 40 images plotted as a function of the numbers of PCs used to parametrize object images. **g**, Example reconstructed images from the three groups defined in **h**. In each pair, the original image is shown on the left, and image reconstructed using neural data are shown on the right. **h**, Distribution of normalized distances between predicted and reconstructed feature vectors. The normalized distance takes account of the fact that the object images used for reconstruction did not include any of the object images shown to the monkey, setting a limit on how good the reconstruction can be (see Methods). A normalized distance of one means that the reconstruction has found the best solution possible. Images were sorted into three groups on the basis of normalized distance. **i**, Distribution of specialization indices SI_{ij} across objects for the NML (left), body (centre) and stubby (right) networks (see Methods and Supplementary Information). Example objects for each network with $\mathrm{SI}_{ij} \approx 1$ are shown. Red bars, objects with SI_{ij} significantly greater than 0 (two-tailed *t*-test, $P < 0.01$).

nature research

Corresponding author(s): Doris Y. Tsao

Last updated by author(s): 02/24/2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Electrophysiological signals were collected by Plexon system (Plexon, Dallas, USA).
Data analysis	Custom code written in MATLAB R2012a , and MATLAB R2018b was used for analysis. Functional imaging data are prcoessed with Freesurfer and FSL.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Custom code used and datasets generated and/or analysed during the current study are available from the corresponding author upon request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes were chosen in a manner commensurate with similar previous studies.
Data exclusions	We recorded single-unit data from every neuron encountered. Only well-isolated units were considered for further analysis; otherwise, every neuron was included for analysis.
Replication	Results were replicated across animals for each experiment.
Randomization	The stimuli were shown in a random order.
Blinding	Investigators were not blinded to experimental groups due to the nature of the experiments.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☐	☒ MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Five male rhesus macaques (Macaca mulatta) of 5-8 years old were used in this study
Wild animals	The study did not involve wild animals.
Field-collected samples	The study did not involve field-collected samples.
Ethics oversight	All procedures conformed to local and US National Institutes of Health guidelines, including the US National Institutes of Health Guide for Care and Use of Laboratory Animals. All experiments were performed with the approval of the Caltech Institutional Animal Care and Use Committee.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Magnetic resonance imaging

Experimental design

Design type	Block design
Design specifications	During the fMRI experiment, stimuli were presented in 24 s blocks at an interstimulus interval of 500 ms
Behavioral performance measures	Monkey's eye position was monitored using an infrared eye tracking system (ISCAN). Juice reward was delivered every 2–4 s if fixation was properly maintained (within a 3.4 degree square window).

Acquisition

Imaging type(s)	Functional and anatomical imaging
Field strength	3 Tesla
Sequence & imaging parameters	T1-weighted anatomical volumes were measured with MP-RAGE sequence(TR 2,300 ms; IR 1,100 ms; TE 3.37 ms; 0.5 mm isotropic voxels) . EPI volumes were acquired in an AC88 gradient insert (Siemens) TR was 2000 ms,TE was 17 ms, voxels were 1 × 1 × 1 mm with an no gap between slices. Matrix size was (96, 96, 64) (read [x], phase [y], slice [z]), the field of view was 96 × 96 mm in-plane. Flip angle was 80°.
Area of acquisition	Whole brain

Diffusion MRI ☐ Used ☒ Not used

Preprocessing

Preprocessing software	Surface reconstruction based on anatomical volumes was performed using FreeSurfer (Massachusetts General Hospital) after skull stripping using FSL's Brain Extraction Tool (University of Oxford). After applying these tools, segmentation was further refined manually.Analysis of functional volumes was performed using the FreeSurfer Functional Analysis Stream (Massachusetts General Hospital). Volumes were corrected for motion and undistorted based on acquired field map.
Normalization	To concatenate different scans, each voxel's responses were percentage transformed with 0 mean value
Normalization template	We did not normalize any imaging data into template. All the analysis were done in the single subject's original space.
Noise and artifact removal	We remove the linear or quadratic trends in the timeseries.
Volume censoring	Motion noises were removed by putting the motion parameters as the regressors in the GLM analysis.

Statistical modeling & inference

Model type and settings	The analysis used only first-level analysis.
Effect(s) tested	We ran t-tests between different conditions within each single subject.

Specify type of analysis: ☒ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference (See Eklund et al. 2016)	All the analyses were done using voxel-wise inference.
Correction	We did not apply any multiple-comparison correction in the fMRI imaging analysis. We set p value at 0.001.

Models & analysis

n/a	Involved in the study
☒	☐ Functional and/or effective connectivity
☒	☐ Graph analysis
☒	☐ Multivariate modeling or predictive analysis

A discrete neuronal circuit induces a hibernation-like state in rodents

https://doi.org/10.1038/s41586-020-2163-6

Received: 27 June 2019

Accepted: 4 March 2020

Tohru M. Takahashi[1,2], Genshiro A. Sunagawa[3]✉, Shingo Soya[2], Manabu Abe[4,5], Katsuyasu Sakurai[2], Kiyomi Ishikawa[3], Masashi Yanagisawa[2], Hiroshi Hama[6], Emi Hasegawa[2], Atsushi Miyawaki[6], Kenji Sakimura[5], Masayo Takahashi[3] & Takeshi Sakurai[1,2,7]✉

Hibernating mammals actively lower their body temperature to reduce energy expenditure when facing food scarcity[1]. This ability to induce a hypometabolic state has evoked great interest owing to its potential medical benefits[2,3]. Here we show that a hypothalamic neuronal circuit in rodents induces a long-lasting hypothermic and hypometabolic state similar to hibernation. In this state, although body temperature and levels of oxygen consumption are kept very low, the ability to regulate metabolism still remains functional, as in hibernation[4]. There was no obvious damage to tissues and organs or abnormalities in behaviour after recovery from this state. Our findings could enable the development of a method to induce a hibernation-like state, which would have potential applications in non-hibernating mammalian species including humans.

Thermostatic animals expend a lot of energy for heat production to maintain their body temperature within a narrow range that is usually higher than the ambient temperature. Some mammals, however, actively lower their body temperature for energy conservation to survive food scarcity in winter—a state known as hibernation[1]. Laboratory mice (*Mus musculus*) do not hibernate, but they exhibit a short-term (less than 24-h) hypometabolic state known as daily torpor[5], during which reducing basal metabolism would be beneficial. Although several experiments have established that both daily torpor and hibernation are regulated by the central nervous system[4,6,7], the mechanisms involved remain unknown. The artificial induction of a hibernation-like hypometabolic state in non-hibernating animals, including humans, would be beneficial for many medical applications[2,3], as well as being of relevance to the possibility of long-distance space exploration in the future[8,9].

Induction of hypometabolism by Q neurons

A hypothalamic neuropeptide, pyroglutamylated RFamide peptide (QRFP), was originally identified using a bioinformatics approach[10,11] and reverse pharmacology[12]. Expression of *Qrfp* mRNA was found to be exclusively localized in the hypothalamus, in which it was distributed in the lateral hypothalamic area (LHA), tuber cinereum and the periventricular nucleus[12]. QRFP has previously been implicated in food intake, sympathetic regulation and anxiety[12,13]. We examined the function of neurons that produce QRFP using mice in which QRFP-producing neurons specifically express iCre (*Qrfp*^{iCre} mice) (Extended Data Fig. 1). We found that excitation of these neurons using a DREADDs (designer receptors exclusively activated by designer drugs) system[14] resulted in a long-lasting decrease in locomotor activity that started almost 30 min after intraperitoneal injection of clozapine-*N*-oxide (CNO) (a DREADD agonist). This effect coincided with a decrease in skin temperature in the interscapular area in which brown adipose tissue (BAT) is located (hereafter referred to as T_{BAT}) (Extended Data Fig. 2). We thus identified *Qrfp* as a genetic marker for hypothermia-inducing neurons. Although QRFP-producing neurons are exclusively located in the hypothalamus, they are distributed among several discrete hypothalamic regions[12] (Extended Data Fig. 1b, c). To identify the regions that were responsible for the effect, we manipulated iCre-positive neurons in these regions individually by injecting Cre-activatable adeno-associated virus (AAV) vectors[15] into the hypothalamus of *Qrfp*^{iCre} mice using several different stereotaxic coordinates. Injection of AAV into an anteromedial region of the hypothalamus resulted in the expression of designated genes such as *GFP* in iCre-positive neurons in the anteroventral periventricular nucleus (AVPe), medial preoptic area (MPA) and periventricular nucleus, but not in the LHA (Fig. 1a). We expressed hM3Dq–mCherry in these regions (that is, in the AVPe, MPA and periventricular nucleus) by injecting *Qrfp*^{iCre} mice with AAV$_{10}$-EF1a-DIO-hM3Dq-mCherry (Q-hM3D mice), and verified by in situ hybridization analysis that mCherry-positive cells expressed *Qrfp* mRNA (Extended Data Fig. 3a). Electrophysiological experiments confirmed that CNO strongly excited these mCherry-positive neurons (Extended Data Fig. 3b–e). We found that intraperitoneal injection of CNO in Q-hM3D mice led to more profound and stable states of hypothermia and immobility than those observed in *Qrfp*^{iCre}*; Rosa26*^{dreaddm3} mice (Fig. 1b, Supplementary Video 1). The hypothermic state, with very low T_{BAT} (below 30 °C), lasted for longer than 48 h. There were many neurons that were double-positive for mCherry and the neuronal activation marker FOS in the AVPe and MPA, confirming the excitation of these neurons in vivo by CNO (Fig. 1c). These observations suggest that iCre-positive neurons around the third ventricle—especially the neurons in the area of the AVPe and MPA (AVPe/MPA)

---

[1]Faculty of Medicine, University of Tsukuba, Tsukuba, Japan. [2]International Institute for Integrative Sleep Medicine (WPI-IIIS), University of Tsukuba, Tsukuba, Japan. [3]Laboratory for Retinal Regeneration, RIKEN Center for Biosystems Dynamics Research, Kobe, Japan. [4]Department of Cellular Neurobiology, Brain Research Institute, Niigata University, Niigata, Japan. [5]Department of Animal Model Development, Brain Research Institute, Niigata University, Niigata, Japan. [6]Laboratory for Cell Function Dynamics, RIKEN Center for Brain Science, Wako, Japan. [7]Life Science Center, Tsukuba Advanced Research Alliance (TARA), University of Tsukuba, Tsukuba, Japan. ✉e-mail: genshiro.sunagawa@riken.jp; sakurai.takeshi.gf@u.tsukuba.ac.jp

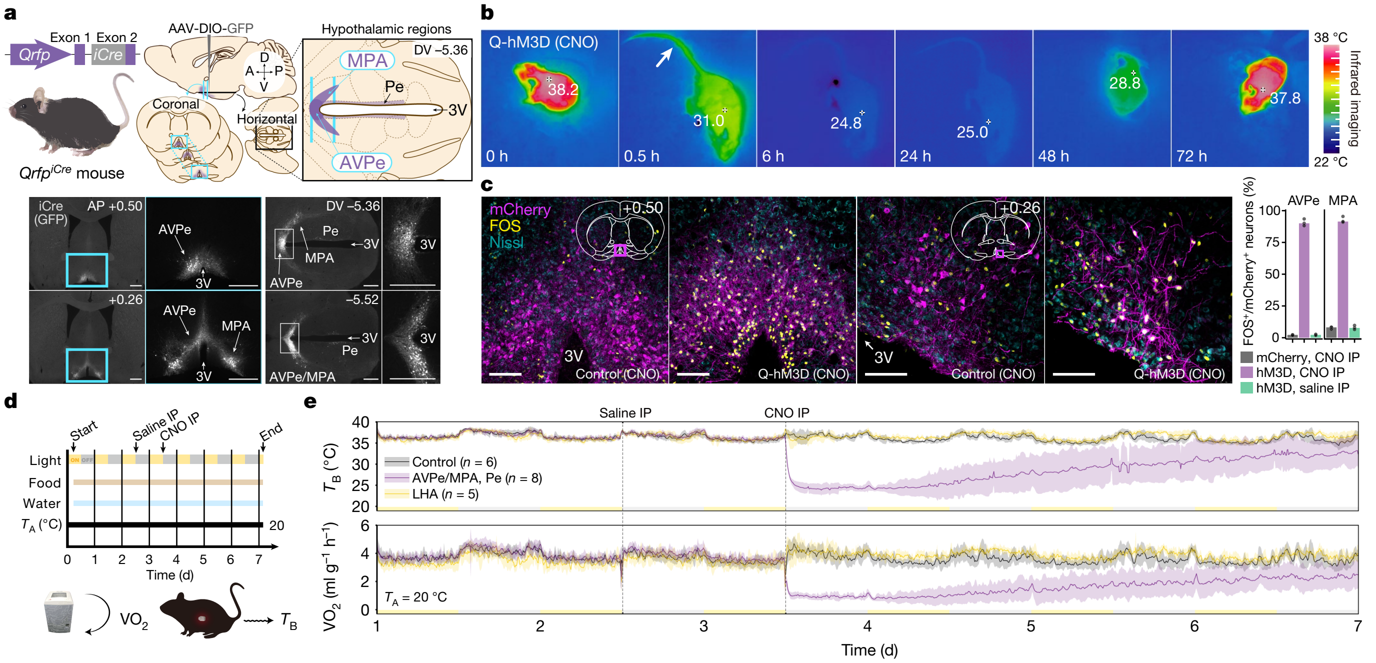


Fig. 1 | Activating *Qrfp*^{iCre} neurons in the hypothalamus lowers body temperature and energy expenditure. a, Distribution of Q neurons in *Qrfp*^{iCre} mice, visualized by GFP expression after injection of AAV$_9$-hSYN-DIO-GFP into the hypothalamic region. A, anterior; P, posterior; D, dorsal; V, ventral; Pe, periventricular nucleus; 3V, third ventricle. The purple fields in the modified stereotaxic brain maps (top) illustrate the locations of Q neurons that are around the third ventricle of the anterior part of the hypothalamus. In the bottom images, the boxed areas on the left are magnified on the right. Scale bars, 500 µm. **b**, Infrared thermal imaging of the surface body temperature of CNO-treated Q-hM3D mice. CNO was injected intraperitoneally at time 0 h. Tail temperature increased at 0.5 h (arrow). The ambient temperature was 23 °C. **c**, Left, brain images immunostained for a neuronal activation marker (FOS) 90 min after intraperitoneal (IP) injection of CNO in control and Q-hM3D mice (AP + 0.50 mm from bregma), and control and Q-hM3D mice (AP + 0.26 mm from bregma). Control mice were *Qrfp*^{iCre} mice that were injected with AAV$_{10}$-EF1α-DIO-mCherry into the AVPe/MPA. Scale bars, 100 µm. Right, bars show the median ratio of FOS-positive neurons in each group and dots represent the raw values of this ratio in each group. **d**, Schematic of metabolic analysis with chemogenetic activation of Q neurons in Q-hM3D mice. Intraperitoneal injection was performed at the beginning of the dark phase. T_A, ambient temperature. Yellow and grey boxes on the *y*-axis show light and dark phases, respectively. **e**, Temporal progression of hypothermia and hypometabolism after DREADD-mediated activation of Q neurons in Q-hM3D mice. T_B, body temperature. Purple, Q-hM3D mice; yellow, *Qrfp*^{iCre} mice with injection of AAV$_{10}$-EF1α-DIO-hM3Dq-mCherry in the LHA (to express hM3Dq in the LHA and tuber cinereum); black, control mice. Line and shading denote mean and s.d. of each group.

(quiescence-inducing neurons or 'Q neurons')—are mainly responsible for the induced hypothermic state.

We next implanted a telemetry temperature sensor in the abdominal cavity of Q-hM3D mice to monitor body temperature, and continuously assessed the metabolism of the mice by analysis of their respiratory gases (Fig. 1d). The results of this experiment confirmed that the CNO-induced hypothermic state in Q-hM3D mice was accompanied by a robust decrease in the rate of oxygen consumption (VO$_2$) (Fig. 1e), and that body temperature decreased concurrently with T_{BAT} after administration of CNO (Extended Data Fig. 4). DREADD-mediated excitation of iCre-positive neurons in the LHA and tuber cinereum in *Qrfp*^{iCre} mice did not induce hypothermia (Fig. 1e), suggesting that iCre-positive cells in lateral regions do not have a role in the effect.

During the state of Q-neuron-induced hypothermia and hypometabolism (QIH), the heart rate of mice decreased considerably, and the respiratory rate was reduced to a level undetectable by the method used, suggesting that the breathing of the mice was shallow (Extended Data Fig. 5a, b). Mice exhibited a very-low-amplitude electroencephalogram (EEG) during QIH, which differed from that observed in sleep (Extended Data Fig. 5b). Serum chemical data showed that blood glucose levels were lower than normal during QIH, presumably owing to decreased gluconeogenesis as a result of low sympathetic activity (Extended Data Fig. 5c).

Although DREADD-mediated effects usually last only a few hours after the injection of CNO[14], DREADD-induced QIH in Q-hM3D mice lasted for several days. At an ambient temperature of 20 °C, QIH (with a body temperature lower than 30 °C) lasted for longer than 48 h after only a single intraperitoneal injection of CNO (1 mg per kg body weight), and it took about a week for VO$_2$ to fully return to normal (Fig. 1e). During QIH, mice showed reduced locomotor activity and food intake, and body weight was lowest at one week after the induction of QIH, followed by a gradual recovery to the normal level (Extended Data Fig. 6a). We performed behavioural tests and found no difference between the group of mice in which QIH was induced and the control group of mice in any of the tests (Extended Data Fig. 6b–d). Gross histological examination of the brain, heart, kidney, liver and muscle did not reveal any damage to tissue after recovery from QIH (Extended Data Fig. 6e). QIH was reproducible in the same mice after repeated injections of CNO (Extended Data Fig. 6f).

Q neurons act on the dorsomedial hypothalamus

After expressing GFP specifically in Q neurons, we observed GFP-positive fibres in several regions of the hypothalamus and brain stem that are implicated in sympathetic regulation and in the control of body temperature[16] (Fig. 2a, Extended Data Fig. 7a–c). We first focused on the dorsomedial hypothalamus (DMH), which receives abundant Q neuron projections (Supplementary Video 2), and in which neurons that promote thermogenesis have previously been identified[17–19]. To examine the function of axonal projections from Q neurons to the DMH, we used an optogenetic approach. We expressed stabilized step function opsin (SSFO)[20] in Q neurons of mice (Q-SSFO mice), and first implanted optic fibres in the AVPe/MPA, where the cell bodies of Q neurons are found (Fig. 2b). Optogenetic excitation of SSFO–eYFP-positive cell bodies by

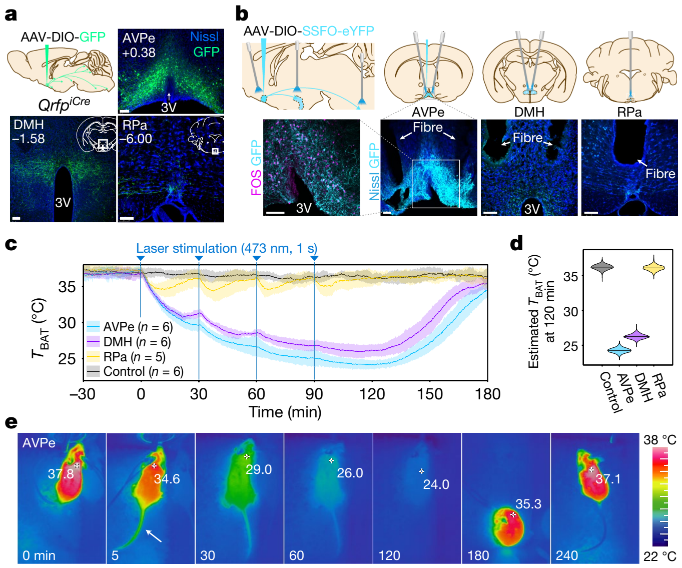


Fig. 2 | Histological and functional analyses of Q neuron projections. **a**, Distribution of cell bodies of Q neurons and a representative axon projection after expression of GFP in the AVPe/MPA. **b**, Strategy for optogenetic excitation of cell bodies or axons of Q neurons in the DMH and RPa. Scale bars, 100 μm (**a**, **b**). **c**, Change in T_{BAT} of Q-SSFO mice during optogenetic excitation of Q neurons. Laser stimulation is shown by blue arrowheads. Line and shading denote mean and s.d. of each group. **d**, The probability density of the estimated T_{BAT} at 30 min after the fourth laser shot. **e**, Representative thermographic images obtained by optogenetic activation of Q neurons (AVPe/MPA). Tail temperature increased at 5 min after the first laser stimulation (arrow).

applying a blue laser (473 nm; one light pulse of 1-s width) rapidly triggered robust hypothermia that lasted for about 30 min (Fig. 2c–e, Supplementary Video 3). Repeating the excitation of Q neurons every 30 min for 2 h resulted in more marked hypothermia, with T_{BAT} dropping to as low as the ambient temperature (22 °C). Many FOS-positive neurons were identified in SSFO–eYFP-positive cells in the AVPe/MPA after excitation (Fig. 2b). Optogenetically induced QIH lasted for less time than QIH that was induced chemogenetically, which suggests that Q neurons are highly sensitive for low levels of CNO or metabotropic signalling mediated by G_q in Q neurons has a role in the long-lasting nature of QIH.

Next, we implanted optic fibres bilaterally in the DMH of Q-SSFO mice and applied optogenetic excitation to the axonal fibres. This manipulation effectively decreased T_{BAT} (Fig. 2c). Optogenetic excitation of Q neuron fibres in the raphe pallidus nucleus (RPa) (Fig. 2b)—a region known to contain sympathetic premotor neurons for thermogenesis through BAT[21]—had subtle effects on T_{BAT} (Fig. 2c). Stimulation of both the cell bodies and the fibres of Q neurons in the DMH or the RPa caused a transient increase in tail temperature, which suggests that peripheral vasodilation is caused by stimulation that acts on both the DMH and the RPa (Extended Data Fig. 7d).

From these results, we postulate that Q neurons act mainly on the DMH (and to a smaller extent on the RPa) to induce QIH. To exclude the possibility that retrograde propagation of axonal excitation by optogenetic stimulation of DMH fibres resulted in the excitation of other collateral projections to induce the effect, we took advantage of the function of SSFO, which is deactivated by yellow light[20]. We found that deactivation of SSFO in the DMH immediately abolished the decrease in T_{BAT} (Extended Data Fig. 7e), further supporting the importance of the Q neuron projections to the DMH.

The thermoregulatory system during QIH

We observed an increase in temperature in the tails of mice immediately after the induction of QIH by either optogenetic or pharmacogenetic excitation of Q neurons. This suggests that peripheral vasodilation was triggered to release heat during the period of decreased body temperature (Figs. 1b, 2e, Extended Data Fig. 7d). The peripheral vasodilation without an increase in body temperature indicates that the reference body temperature (T_R), or the theoretical set-point of body temperature, was reset to a lower value than that in a normal state—a feature of hibernation[4]. To further investigate this possibility, we characterized the thermoregulatory system during QIH. When an animal is doing no external work and has a stable metabolism, the heat conductance (*G*), negative feedback gain of heat production (*H*), and T_R can be estimated from the body temperature and VO_2 at different ambient temperatures[5]. We recorded the body temperature and VO_2 of Q-hM3D mice during QIH at various ambient temperatures (8, 12, 16, 20, 24, 28 and 32 °C) (Fig. 3a). The average body temperature and VO_2 11 h after intraperitoneal injection of saline or CNO were used to estimate the values of *G*, *H* and T_R (Fig. 3b, c). The 89% highest-posterior-density interval (HPDI; hereafter, the 89% HPDI is indicated by two numbers in square brackets) of the heat conductance (*G*) was [0.212, 0.221] $\mathrm{ml\,g^{-1}\,h^{-1}\,°C^{-1}}$ in a normal state and [0.182, 0.220] $\mathrm{ml\,g^{-1}\,h^{-1}\,°C^{-1}}$ in QIH (Fig. 3d–f), suggesting that heat conductance is comparable under normal and QIH conditions. This differs from daily torpor, during which the value of *G* is lower than that observed in normal conditions[5]. For the negative feedback gain of heat production parameter (*H*), the 89% HDPI was [3.43, 8.72] $\mathrm{ml\,g^{-1}\,h^{-1}\,°C^{-1}}$ in a normal state and [0.181, 0.369] in QIH (Fig. 3g–i). This represents a 95.3% reduction in the median value of *H* in the QIH state compared to the normal state, suggesting a robust decrease in heat production. This decrease in *H* resembles the previously reported reduction of *H* during fasting-induced daily torpor (FIT)[5]. Notably, T_R was estimated to be [36.04, 36.60] °C in the normal condition and [26.83, 29.13] °C in QIH (Fig. 3g, j). The difference in the median T_R was 8.41 °C, and the posterior distribution of the difference (ΔT_R) was [7.18, 9.57] °C, demonstrating a reduction in T_R during QIH (Fig. 3k). Considering the very small shift in T_R that is observed in FIT[5], this observation underscores the similarity between QIH and hibernation and the difference between QIH and daily torpor—although we should note that T_R was estimated only in a narrow interval of ambient temperatures here, which is different from estimations in hibernators.

To provide more evidence of the reduction in T_R during QIH, which is a prominent characteristic of hibernation, we observed the relationship between the posture and metabolism of mice when the ambient temperature was changed during QIH (Fig. 3l, Extended Data Fig. 8a, b). Notably, at an ambient temperature of 28 °C, mice showed an extended posture during QIH—a posture that is normally seen in animals exposed to a hot environment (image D in Fig. 3l). This was different from the typical sitting posture that was observed during FIT at an ambient temperature of 28 °C (B in Fig. 3l). These observations further demonstrate that T_R was lower in QIH than in FIT or a normal state. Moreover, when the ambient temperature was lowered to 12 °C, the mouse returned to a sitting posture (E in Fig. 3l), exhibited shivering (Supplementary Video 4) and showed an increase in VO_2. These results support the hypothesis that during QIH, T_R is lowered, but bodily functions and behaviour are still regulated to adapt to a change in ambient temperature.

During QIH, mice showed a higher VO_2 when they were exposed to an ambient temperature below 16 °C compared with a temperature of 20 °C or 28 °C (Fig. 3b, Extended Data Fig. 8c). This shows the similarity between QIH and hibernation, in which mice showed an increased metabolism when the ambient temperature was lowered to a certain level[22]. This regulated hypometabolic feature of QIH was also confirmed in individual mice (Fig. 3l, Extended Data Fig. 8a, b). In addition, the respiratory quotient dropped to a level close to 0.7 at all ambient temperatures during QIH, implying that the energy source shifted from carbohydrates to lipids (Fig. 3l, Extended Data Fig. 8a). This agrees with the reduction in respiratory quotient that has been reported previously in deep torpor during hibernation[23]. The behavioural and metabolic responses of mice during QIH were quite different from

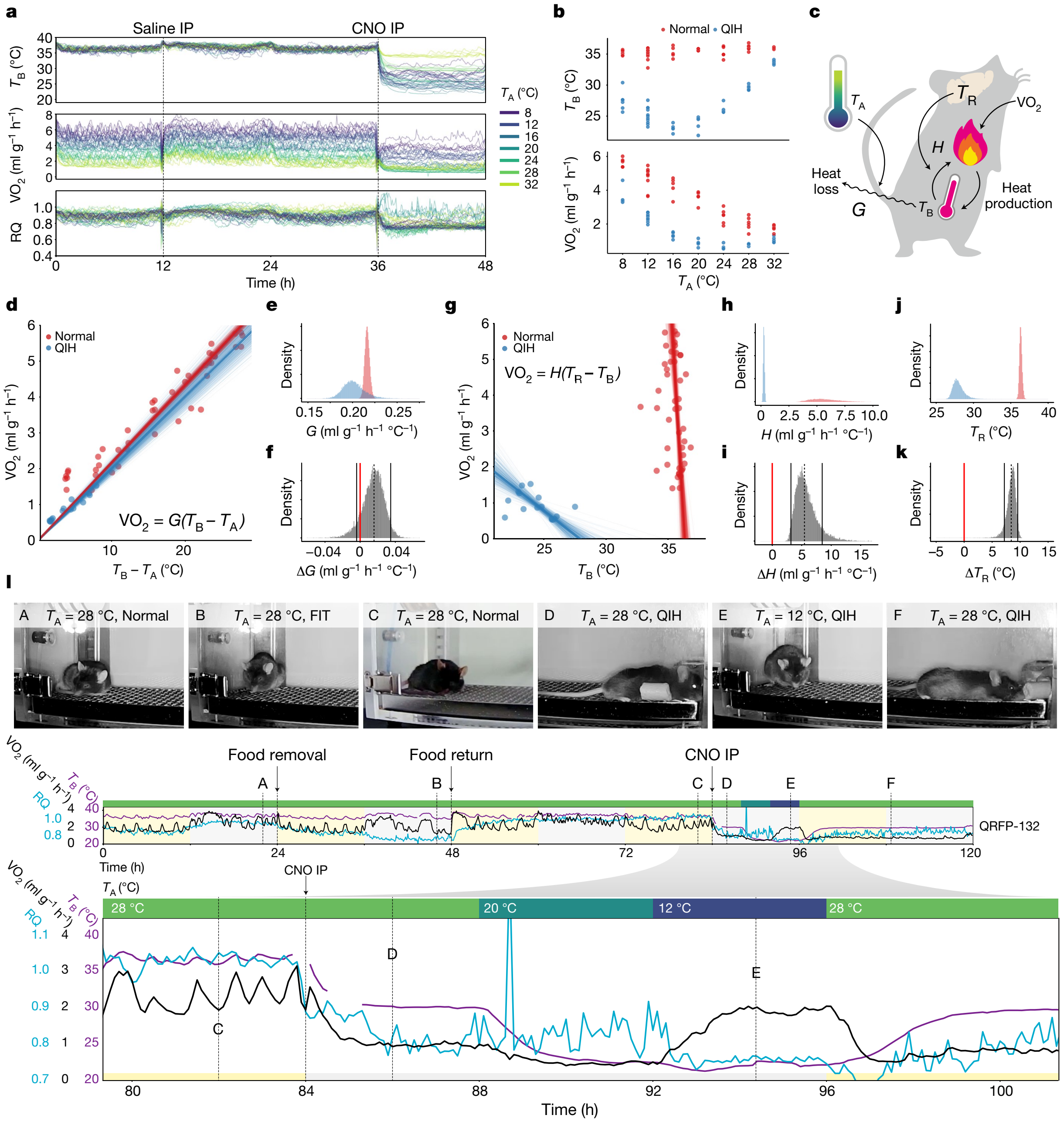


Fig. 3 | Q-neuron-induced hypometabolism is accompanied by a lowered set-point of body temperature. **a**, Change in body temperature, VO_2 and respiratory quotient (RQ) during QIH at various ambient temperatures. Each line denotes one mouse. **b**, Minimum body temperature (top) and VO_2 (bottom) under normal and QIH conditions. **c**, Schematic of heat-production and heat-loss pathways in mice. Heat loss is proportional to the difference between T_A and T_B at factor G. Heat production is governed by the difference between T_R and T_B at factor H. **d**, Relationship between $T_B - T_A$ and VO_2 at various values of T_A. The slope of the curve denotes G. Dots are recorded data, thick lines are drawn from the median of posterior G and thin lines are drawn from 500 randomly selected values of G from posterior samples. **e**, Posterior distribution of estimated G. **f**, Difference in G from QIH to the normal condition. **g**, Relationship between T_B and VO_2 at various values of T_A. The negative slope of the curves denotes H and the x-axis intercept denotes T_R. Dots and lines as in **d**. **h**, Posterior distribution of estimated H. **i**, Difference in H from QIH to the normal condition. **j**, Distribution of estimated T_R. **k**, Difference in T_R from QIH to the normal condition. **l**, Metabolic transition and postures during QIH within an individual mouse. The bottom chart is the timewise magnification of the top chart. The mouse shows a curled-up posture during FIT at T_A = 28 °C (B), but an extended posture during QIH at T_A = 28 °C (D). Even during QIH when T_A is lowered to 12 °C, the mouse assumes a curled-up posture, as in FIT (E), indicating that it is avoiding heat loss. During QIH, the respiratory quotient always decreases to a level close to 0.7, independent of T_A—a feature that is shared with hibernation. Three other examples are shown in Extended Data Fig. 8a.

those observed during a normal state, in which the primary function of the thermoregulatory system is to maintain the body temperature within a narrow range (Extended Data Fig. 8d). QIH is also completely different from an anaesthetized state, in which mice showed neither an increase of VO_2 nor a change in posture when exposed to low ambient temperature (Extended Data Fig. 9).

Neurotransmission in Q neurons

Expression of tetanus toxin light chain (TeTxLC) in mouse Q neurons (Q-TeTxLC mice) completely abolished the induction of QIH (Extended Data Fig. 10a, b), suggesting that SNARE-mediated neurotransmission in Q neurons is necessary for the induction of QIH. To ascertain

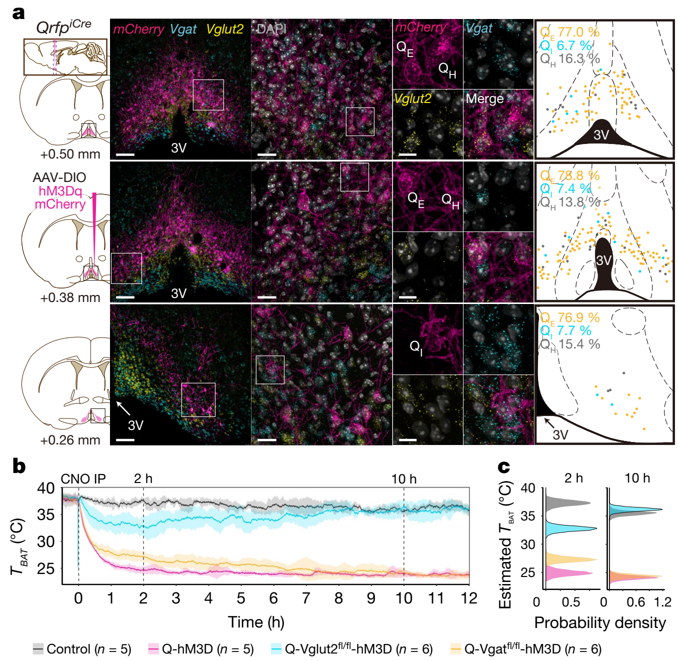


Fig. 4 | Glutamatergic and GABAergic neurotransmission of Q neurons are both involved in inducing QIH. a, In situ hybridization analysis showing Q neurons expressing *Vgat* and/or *Vglut2* in Q-hM3D mice. The left column shows schematics of coronal brain sections with positions from bregma. The middle images show the representative distribution of Q_E, Q_I and Q_H neurons. The right column shows the distribution of these neurons. We calculated the percentages of neurons in *mCherry*-expressing cells (counted on each coordinate from two slices prepared from independent Q-hM3D mouse brains) that were positive for *Vglut2* (Q_E; 315 out of 404 cells), *Vgat* (Q_I; 29 out of 404 cells) and both *Vglut2* and *Vgat* (Q_H; 60 out of 404 cells). Scale bars, 100 μm (left); 25 μm (middle); 10 μm (right). **b**, The absence of *Vgat* or *Vglut2* affects QIH. In Q-Vgat$^{fl/fl}$-hM3D mice (orange line), CNO injection effectively induced QIH, and T_{BAT} reached a level comparable to that observed during QIH in Q-hM3D mice (magenta line). In Q-Vglut2$^{fl/fl}$-hM3D mice (cyan line), the hypothermic effect of CNO injection was smaller and shorter. Line and shading denote mean and s.d. of each group. **c**, Estimated T_{BAT} of each genotype at 2 h and 10 h after CNO injection. Of note, T_{BAT} in Q-Vgat$^{fl/fl}$-hM3D mice showed little difference compared to that in Q-hM3D mice at 10 h after CNO injection, suggesting that glutamatergic neurotransmission from Q neurons is indispensable for maintaining a hypometabolic state in QIH.

whether Q neurons are inhibitory or excitatory, we examined the colocalization of *mCherry* expression with that of the genes that encode vesicular glutamate transporter 2 (VGLUT2) (*Vglut2*; also known as *Slc17a6*) and vesicular GABA transporter (VGAT) (*Vgat*; also known as *Slc32a1*) in Q-hM3D mice. We found that there are at least three populations of Q neurons: (i) Q_E (excitatory) neurons that are positive for *Vglut2* (77.9%); (ii) Q_I (inhibitory) neurons that are positive for *Vgat* (7.2%); and (iii) Q_H (hybrid) neurons that are positive for both *Vglut2* and *Vgat* (14.9%) (Fig. 4a). The proportions of these three populations were similar in all selected regions, and the neurons were intermingled with one another. Q_E neurons constitute the largest population, which is consistent with previous single-cell transcriptome studies of neurons of the preoptic area: in one report[24], 12 out of 31,299 cells were *Qrfp*-positive, and there were 7 excitatory and 4 inhibitory cells; and in another study[25], 16 out of 14,437 cells from the hypothalamus were *Qrfp*-positive, and there were 5 excitatory cells and 1 inhibitory cell.

Next, we mated *Qrfp*^{iCre} mice with *Slc32a1*$^{fl/fl}$ or *Slc17a6*$^{fl/fl}$ mice to obtain *Qrfp*^{iCre} mice that lack the expression of VGAT or VGLUT2 in Q neurons (which we term Q-Vgat$^{fl/fl}$ and Q-Vglut2$^{fl/fl}$ mice, respectively). After injecting AAV$_{10}$-EF1a-DIO-hM3Dq-mCherry into the AVPe/MPA (Q-Vgat$^{fl/fl}$-hM3D and Q-Vglut2$^{fl/fl}$-hM3D mice), we examined how the absence of VGAT or VGLUT2 affects QIH. In Q-Vgat$^{fl/fl}$-hM3D mice, injection of CNO effectively induced QIH, and T_{BAT} reached a level comparable to that observed during QIH in Q-hM3D mice (Fig. 4b, c). However, the initial reduction in T_{BAT} after injection of CNO was notably slower in these mice compared with that in control Q-hM3D mice. In Q-Vglut2$^{fl/fl}$-hM3D mice, although CNO injection induced a decrease in T_{BAT}, the effect was smaller and shorter than that in Q-hM3D mice. These results suggest that both glutamatergic and GABAergic neurotransmission cooperatively induce QIH.

Discussion

We have demonstrated that a hypothalamic population of neurons with a particular genetic identity (that is, expressing *Qrfp*) and spatial location (that is, within the AVPe/MPA) exists in mice, and that chemogenetic excitation of this population induces QIH—an extremely long-lasting state of regulated hypometabolism. QIH shares four key properties with hibernation. First, the hypothermic and hypometabolic state lasts for more than 24 h; second, the theoretical T_R is lowered but the thermoregulatory system remains functional—even under hypothermic conditions that are possibly harmful—to adapt to changes in the outer environment[4]; third, despite physiological functions being suppressed, with mice showing a slow heart rate, weak respiration, and low-voltage EEG, there is no tissue damage—a signature feature of hibernation[26]; and finally, mice spontaneously recover from QIH without any external manipulation.

By analysing the expression of FOS, a previous study showed that cells near the third ventricle are activated during hibernation in the thirteen-lined ground squirrel[27]. This activation pattern is similar to the region in which Q neurons are localized, suggesting that hibernators might also use Q neurons to induce hibernation. As the ability to hibernate exists among distantly related mammals—including rodents, the caniformia and even primates[28]—it is reasonable to hypothesize that the neuronal mechanism of hibernation is preserved among a broad range of mammalian species, although the system is not mobilized in non-hibernating species.

We identified the DMH as the major effector site of Q neurons. Glutamatergic DMH neurons located in the dorsal area of the DMH that send projections to the RPa have previously been shown to modulate BAT thermogenesis, which suggests that there is a connection between the DMH and the RPa[18,19]. A previous report showed that suspended animation in rats was induced by microinjection of muscimol into the RPa, which was accompanied by initial vasodilation in the tail[29]. These observations indicate that GABAergic neurotransmission by Q_I neurons might inhibit the circuit between glutamatergic DMH neurons and the RPa to induce tail vasodilation. Q_E neurons might excite another subset of DMH neurons to inhibit heat production. Future studies that identify QIH-inducing neurons in the DMH will enable the mechanism of QIH to be further elucidated.

The physiological role of Q neurons remains unknown. One possible function in which they are involved is the regulation of daily torpor. Although QIH is more similar to hibernation than to daily torpor, a shared mechanism might be present between hibernation and daily torpor[30]. In line with this, we found that the normal architecture of FIT was disrupted when neurotransmission of Q neurons was blocked in Q-TeTxLC mice (Extended Data Fig. 10c–f), suggesting that the function of Q neurons is necessary to evoke the rapid decrease of body temperature that occurs during FIT—although QRFP did not itself have a role in regulating FIT (Extended Data Fig. 10g). In addition, Q-TeTxLC mice showed less circadian fluctuation in body temperature than did control mice, which suggests that Q neurons might be involved in the circadian regulation of body temperature (Extended Data Fig. 10h). Q neurons receive input from the preoptic area and paraventricular hypothalamic nucleus (Extended Data Fig. 11a–c), indicating that they could receive circadian information from these regions of the brain. Because Q neurons are localized along the third ventricle and their dendrites extend along the ependyma of the third ventricle and regions of nearby circumventricular organs (Fig. 1a), they might also sense humoral factors that are released by tanycytes and ependymal cells[31], or by factors in the cerebrospinal fluid. Our histological study also suggests that many Q neurons constitute a unique subpopulation of previously shown warm-sensitive neurons that co-express the neuropeptides BDNF and PACAP (BDNF/PACAP neurons) in the preoptic area[32] (Extended Data Fig. 11d, e).

We found that, notably, mice can enter a hibernation-like multi-day state of torpor by stimulating a defined neuronal population. Moreover, we observed that excitation of AVPe/MPA neurons, including Q neurons, also induced a QIH-like hypometabolic state in rats—a species that shows neither hibernation nor daily torpor (Extended Data Fig. 12). This induction of a hibernation-like condition in a non-hibernating mammal is a step forward in our understanding of the neuronal mechanisms of regulated hypometabolism, and will enable further investigation into how each tissue adopts a hibernation-like state. Furthermore, the future development of a method that enables the selective manipulation of Q neurons could provide a new approach through which a QIH-like state of synthetic hibernation could be induced in humans. This would have many potential clinical applications, including the reduction of systemic tissue damage following heart attacks or strokes, and the preservation of organs for transplants.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2163-6.

1. Geiser, F. Hibernation. *Curr. Biol*. **23**, R188–R193 (2013).
2. Melvin, R. G. & Andrews, M. T. Torpor induction in mammals: recent discoveries fueling new ideas. *Trends Endocrinol. Metab*. **20**, 490–498 (2009).
3. Bouma, H. R. et al. Induction of torpor: mimicking natural metabolic suppression for biomedical applications. *J. Cell. Physiol*. **227**, 1285–1290 (2012).
4. Jastroch, M. et al. Seasonal control of mammalian energy balance: recent advances in the understanding of daily torpor and hibernation. *J. Neuroendocrinol*. **28**, 12347 (2016).
5. Sunagawa, G. A. & Takahashi, M. Hypometabolism during daily torpor in mice is dominated by reduction in the sensitivity of the thermoregulatory system. *Sci. Rep*. **6**, 37011 (2016).
6. Vicent, M. A., Borre, E. D. & Swoap, S. J. Central activation of the A_1 adenosine receptor in fed mice recapitulates only some of the attributes of daily torpor. *J. Comp. Physiol. B* **187**, 835–845 (2017).
7. Hitrec, T. et al. Neural control of fasting-induced torpor in mice. *Sci. Rep*. **9**, 15462 (2019).
8. Griko, Y. & Regan, M. D. Synthetic torpor: a method for safely and practically transporting experimental animals aboard spaceflight missions to deep space. *Life Sci. Space Res*. **16**, 101–107 (2018).
9. Choukèr, A., Bereiter-Hahn, J., Singer, D. & Heldmaier, G. Hibernating astronauts—science or fiction? *Pflugers Arch*. **471**, 819–828 (2019).
10. Fukusumi, S. et al. A new peptidic ligand and its receptor regulating adrenal function in rats. *J. Biol. Chem*. **278**, 46387–46395 (2003).
11. Chartrel, N. et al. Identification of 26RFa, a hypothalamic neuropeptide of the RFamide peptide family with orexigenic activity. *Proc. Natl Acad. Sci. USA* **100**, 15247–15252 (2003).
12. Takayasu, S. et al. A neuropeptide ligand of the G protein-coupled receptor GPR103 regulates feeding, behavioral arousal, and blood pressure in mice. *Proc. Natl Acad. Sci. USA* **103**, 7438–7443 (2006).
13. Okamoto, K. et al. QRFP-deficient mice are hypophagic, lean, hypoactive and exhibit increased anxiety-like behavior. *PLoS ONE* **11**, e0164716 (2016).
14. Roth, B. L. DREADDs for neuroscientists. *Neuron* **89**, 683–694 (2016).
15. Atasoy, D., Aponte, Y., Su, H. H. & Sternson, S. M. A FLEX switch targets channelrhodopsin-2 to multiple cell types for imaging and long-range circuit mapping. *J. Neurosci*. **28**, 7025–7030 (2008).
16. Nakamura, K. Central circuitries for body temperature regulation and fever. *Am. J. Physiol. Regul. Integr. Comp. Physiol*. **301**, R1207–R1228 (2011).
17. Zhao, Z.-D. et al. A hypothalamic circuit that controls body temperature. *Proc. Natl Acad. Sci. USA* **114**, 2042–2047 (2017).
18. Machado, N. L. S. et al. A glutamatergic hypothalamomedullary circuit mediates thermogenesis, but not heat conservation, during stress-induced hyperthermia. *Curr. Biol*. **28**, 2291–2301 (2018).
19. Piñol, R. A. et al. *Brs3* neurons in the mouse dorsomedial hypothalamus regulate body temperature, energy expenditure, and heart rate, but not food intake. *Nat. Neurosci*. **21**, 1530–1540 (2018).
20. Yizhar, O. et al. Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature* **477**, 171–178 (2011).
21. Morrison, S. F. Central control of body temperature. *F1000Res*. **5**, 880 (2016).
22. Ortmann, S. & Heldmaier, G. Regulation of body temperature and energy requirements of hibernating alpine marmots (*Marmota marmota*). *Am. J. Physiol. Regul. Integr. Comp. Physiol*. **278**, R698–R704 (2000).
23. Snapp, B. D. & Heller, H. C. Suppression of metabolism during hibernation in ground squirrels (*Citellus lateralis*). *Physiol. Zool*. **54**, 297–307 (1981).
24. Moffitt, J. R. et al. Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science* **362**, eaau5324 (2018).
25. Chen, R., Wu, X., Jiang, L. & Zhang, Y. Single-cell RNA-seq reveals hypothalamic cell diversity. *Cell Rep*. **18**, 3227–3241 (2017).
26. Walker, J. M., Glotzbach, S. F., Berger, R. J. & Heller, H. C. Sleep and hibernation in ground squirrels (*Citellus* spp): electrophysiological observations. *Am. J. Physiol. Integr. Comp. Physiol*. **233**, R213–R221 (1977).
27. Bratincsák, A. et al. Spatial and temporal activation of brain regions in hibernation: *c-fos* expression during the hibernation bout in thirteen-lined ground squirrel. *J. Comp. Neurol*. **505**, 443–458 (2007).
28. Dausmann, K. H., Glos, J., Ganzhorn, J. U. & Heldmaier, G. Physiology: hibernation in a tropical primate. *Nature* **429**, 825–826 (2004).
29. Cerri, M. et al. The inhibition of neurons in the central nervous pathways for thermoregulatory cold defense induces a suspended animation state in the rat. *J. Neurosci*. **33**, 2984–2993 (2013).
30. Heldmaier, G., Ortmann, S. & Elvert, R. Natural hypometabolism during hibernation and daily torpor in mammals. *Respir. Physiol. Neurobiol*. **141**, 317–329 (2004).
31. Oomura, Y. et al. A new brain glucosensor and its physiological significance. *Am. J. Clin. Nutr*. **55**, 278S–282S (1992).
32. Tan, C. L. et al. Warm-sensitive neurons that control body temperature. *Cell* **167**, 47–59 (2016).

Methods

Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized, and the investigators were not blinded to allocation during experiments and outcome assessment.

Animals

All animal experiments were performed at the International Institute of Integrative Sleep Medicine (IIIS), Tsukuba University and RIKEN Center for Biosystems Dynamics Research (BDR), according to their guidelines for animal experiments. They were approved by the animal experiment committees of each institute, and thus were in accordance with NIH guidelines. Except during torpor-inducing experiments, animals were given food and water ad libitum and maintained at an ambient temperature of 23 °C at IIIS and 22 °C at BDR and a relative humidity of 50%, with a 12-h light/12-h dark cycle. Because we found that mice that weighed more than 34 g did not reproducibly exhibit FIT, we excluded mice heavier than 34 g in daily torpor experiments.

Qrfp^{iCre} mice were generated by homologous recombination in C57BL/6N embryonic stem cells and implantation in 8-cell-stage embryos (ICR). A targeting vector was designed to replace the entire coding region of the *prepro-Qrfp* sequence in exon 2 of the *Qrfp* gene with *iCre* and a pgk-Neo cassette so that the endogenous *Qrfp* promoter drives expression of *iCre* (Extended Data Fig. 1). Chimeric mice were crossed with C57BL/6J females (Jackson Laboratory). The Pgk-Neo cassette was deleted by crossing the mice with FLP66 mice, which had been backcrossed to C57BL/6J mice at least 10 times. Initially, F$_1$ hybrids from mating heterozygotes with heterozygotes were generated. We backcrossed them to C57BL/6J mice at least 8 times. All experiments were performed on *iCre* heterozygotes, unless indicated otherwise. *Rosa26*^{dreaddm3} and *Rosa26*^{dreaddm4} mice were generated by homologous recombination in C57BL/6N embryonic stem cells, followed by the same procedure as in *Qrfp*^{iCre} mice, as described above. Targeting vectors are shown in Extended Data Fig. 2a. *Slc32a1*^{tm1Lowl} (referred to as *Vgat*$^{fl/fl}$) mice and *Slc17a6*^{tm1Lowl} (*Vglut2*$^{fl/fl}$) mice were obtained from the Jackson Laboratory (stock no. 012897 and 012898, respectively). Wistar rats were purchased from Oriental Yeast Co.

Viruses

AAVs were produced using a triple-transfection, helper-free method as previously described[33]. The final purified viruses were stored at −80 °C. Titres of recombinant AAV vectors were determined by quantitative PCR: AAV$_{10}$-EF1α-DIO-TVA-mCherry, 4×10^{13}; AAV$_{10}$-CAG-DIO-RG, 1×10^{13}; AAV$_{10}$-EF1α-DIO-hM3Dq-mCherry, 1.64×10^{12}; AAV$_{10}$-EF1α-DIO-mCherry, 1.44×10^{12}; AAV$_{10}$-EF1α-DIO-SSFO-EYFP, 1.35×10^{12}; AAV$_9$-SYN-DIO-TeTxLC-GFP, 6.24×10^{14}; AAV$_9$-hSYN-DIO-GFP, 4×10^{12}; AAV$_{10}$-CaMKIIα-hM3Dq-mCherry, 1.4×10^{14}; AAV$_9$-hSYN-DIO-GCaMP6 s, 1×10^{15} genome copies per ml. Recombinant rabies vectors were produced by a previously reported procedure[34,35]. The titre of SADΔG-GFP(EnvA) was 4.2×10^8 infectious units per ml.

Surgery

For injection of AAV vectors, male *Qrfp*^{iCre} heterozygous mice (8–12 weeks old) and male Wistar rats (8 weeks old) were anaesthetized with isoflurane and positioned in a stereotaxic frame (David Kopf Instruments). Virus was delivered into the target site at a controlled rate of 0.1 μl per min using a Hamilton needle syringe. The needle was kept in place for 10 min after injection. The waiting period for recovery and virus expression for the experiments was at least 2 weeks except as noted.

For the chemogenetic manipulation in Fig. 1, *Qrfp*^{iCre} mice underwent injection of AAV$_{10}$-EF1α-DIO-hM3Dq-mCherry into the hypothalamus (to express in the AVPe/MPA: anterior–posterior (AP), −0.22 mm; medial–lateral (ML), ± 0.25 mm; dorsal–ventral (DV), −5.50 mm; 0.50 μl in each site; to express in the LHA: AP, −1.00 mm; ML, ± 1.00 mm; DV, −5.00 mm; 0.30 μl in each site).

For the identification of the distribution and the axonal projections of Q neurons in Figs. 1, 2, we injected 0.30 μl AAV$_9$-hSYN-DIO-GFP into the hypothalamus (AVPe/MPA: AP, −0.22 mm; ML, 0.25 mm; DV, −5.50 mm) unilaterally.

For the optogenetic manipulations in Fig. 2 and Extended Data Fig. 7, we injected 0.3 μl AAV$_{10}$-EF1α-DIO-SSFO-EYFP into the AVPe/MPA (AP, 0.38 mm; ML, 0.25 mm; DV, −5.25 mm from bregma) unilaterally. Optical fibres were then implanted bilaterally above the AVPe/MPA (AP, 0.38 mm; ML, ± 0.25 mm; DV, −5.00 mm), bilaterally above the DMH (AP, −1.70 mm; ML, ± 0.25 mm; DV, −4.75 mm) or unilaterally above the RPa (AP, −6.00 mm; ML, 0.00 mm; DV, −5.50 mm). After a recovery period of at least three weeks in individual cages after injection, mice were subjected to infrared thermal-imaging experiments. Behavioural data were only included if these viruses were targeted specifically to Q neurons and the fibre-optic implants were precisely placed.

For demonstrating the identity and detailed anatomical location of Q neurons (Fig. 4a, Extended Data Fig. 11) and chemogenetic manipulation of conditional knockout mice (Fig. 4b), we injected 0.20 μl AAV$_{10}$-EF1a-DIO-hM3Dq-mCherry into the AVPe/MPA (AP, 0.38 mm; ML, 0.25 mm; DV, −5.25 mm from bregma) unilaterally.

For silencing experiments, we injected mixed AAV (0.20 μl AAV$_9$-SYN-DIO-TeTxLC-GFP and 0.20 μl AAV$_{10}$-EF1α-DIO-hM3Dq-mCherry) into the AVPe/MPA and periventricular nucleus region (AP, −0.22 mm; ML, ± 0.25 mm; DV, −5.50 mm) bilaterally (Extended Data Fig. 10a, b), and we injected 0.30 μl AAV$_9$-SYN-DIO-TeTxLC-GFP into the same injection sites (Extended Data Fig. 10c–g).

For rat experiments in Extended Data Fig. 12, we injected 0.20 μl AAV$_{10}$-CaMKIIα-hM3Dq-mCherry into the AVPe/MPA (AP, 0.12 mm; ML, ± 0.40 mm; DV, −8.50 mm from bregma) bilaterally.

Drug administration

CNO (Abcam, ab141704) was dissolved in normal saline at a dose of 100 μg ml^{-1} and frozen at −20 °C. The CNO solution was thawed on site, and it was administered intraperitoneally at a dose of 1 mg kg^{-1} for mice and 5 mg kg^{-1} for rats.

Biological signal recordings

For thermographic analysis, mice were put in experimental cages (25 × 15 × 10 cm) and monitored using an infrared thermal-imaging camera (InfReC R500EX, Nippon Avionics) positioned 30 cm above the cage floor. To clearly detect surface temperature, the back hair was removed with hair clippers one day before starting the experiment. Thermograms of DREADD and optogenetic experiments were collected at 0.5 Hz and 1 Hz, respectively and analysed with InfReC Analyzer NS9500 Professional software (Nippon Avionics). The highest temperature in one frame was used as the T_{BAT} of the mouse (Figs. 1b, 2c–e, 4b, c, Extended Data Figs. 4, 7d, e). These experiments were performed in a temperature-controlled chamber (HC-100, Shin Factory) at 22 °C.

For recordings of core body temperature, VO$_2$, EEG, electrocardiogram (ECG) and respiratory pattern, each animal was housed in a temperature-controlled chamber (HC-100, Shin Factory or LP-400P-AR, Nippon Medical & Chemical Instruments). To record the body temperature continuously, a telemetry temperature sensor (TA11TA-F10, DSI) was implanted in the animal's abdominal cavity under general inhalation anaesthesia at least seven days before recording. Artefacts on the recording of body temperature that were caused by animal movements were filtered by a custom R script based on a secondary trend model interpolation. VO$_2$ and the carbon dioxide output rate (VCO$_2$) of the animal were continuously recorded with a respiratory-gas analyser (ARCO-2000 mass spectrometer, ARCO system). The respiratory quotient was calculated as the ratio of VCO$_2$ to VO$_2$.

EEG and ECG were recorded by implanted telemetry transmitters (F20-EET or HD-X02, DSI). For EEG recording, two stainless-steel screws

(1-mm diameter) were soldered to the wires of telemetry transmitters and inserted through the skull of the cortex (AP, 1.00 mm; right, 1.50 mm from bregma or lambda) under general anaesthesia. Two other wires from the transmitter were placed on the surface of the thoracic cavity to record ECG. Animals were allowed at least 10 days to recover from surgery. The EEG and ECG data-collecting system consisted of transmitters, an analogue–digital converter and a recording computer with the software Ponemah Physiology Platform (v.6.30, DSI). The sampling rate was 500 Hz for both EEG and ECG, and data were converted to ASCII format for review. Heart rate was detected by QRS-complex peak-detection analysis of the EEG. Analysis of the heart-rate variability (HRV) was performed by resampling the train of R–R-interval (ms) into a 20-Hz time series followed by power-spectrum analysis by a 60-s window. The R–R-interval time series was detrended and a Hanning window was applied before frequency-domain analysis. High-frequency and low-frequency power (ms^2) were defined as the sum of 1.5–4 Hz and 0.4–1.5 Hz of the power spectrum of HRV, respectively[36].

Respiratory flow was recorded by a non-invasive respiratory flow recording system[37]. In brief, mice were placed in a metabolic chamber (TMC-1213-PMMA, Minamiderika Shokai), which had airflow of at least 0.3 l min^{-1}. The chamber was connected to a pressure sensor (PMD-8203-3G, Biotex), which detected the pressure difference between the outside and inside of the chamber. When the mouse is breathing, the pressure difference from outside to inside becomes larger during inspiration and smaller on expiration[37]. The analogue signal output from the sensor was digitized at 250 Hz by an AD converter (NI-9205, National Instruments) and stored on a computer by data-logging software developed by Biotex.

Metabolism recording during general anaesthesia

In addition to the body temperature, VO_2 and video recording described above (see 'Biological signal recordings'), the inlet of the metabolic chamber was directly connected to the outlet of the inhalation anaesthesia machine (NARCOBIT-E, Natsume Seisakusho). Animals were given 1% isoflurane at an ambient temperature of 28 °C for 30 min, followed by 90 min at an ambient temperature of 12 °C. After the experiment, the animals were warmed on a hot plate and recovery was confirmed.

Chemical analysis of blood

Blood was collected from anaesthetized mice by left ventricular puncture using a 25-gauge needle. The collected blood was stored on ice no longer than 2 h. Samples were centrifuged at 2,000*g* for 10 min at 4 °C, and supernatants were collected and frozen at −30 °C. Frozen serum samples were sent to Fujifilm Wako Pure Chemical Corporation to measure Na^+ (mM), K^+ (mM), Cl^- (mM), aspartate aminotransferase (IU l^{-1}), alanine transaminase (IU l^{-1}), lactic acid dehydrogenase (IU l^{-1}), creatine kinase (IU l^{-1}), glucose (mg dl^{-1}) and total serum ketone (μmol l^{-1}) levels.

Immunohistochemistry

Animals were deeply anaesthetized with isoflurane. They were perfused transcardially with 10% sucrose in water, followed by ice-cold 4% paraformaldehyde in 0.1 M phosphate buffer pH 7.4 (4% PFA), and the brain was removed. Brains were post-fixed overnight in 4% PFA at 4 °C, incubated overnight in 30% sucrose in 0.1 M phosphate buffered saline pH 7.4 (PBS) at 4 °C, immersed in Tissue-Tek OCT compound (Sakura) in cryomolds and frozen at −80 °C until sectioning. Brains were sliced coronally using a cryostat (CM1860, Leica) every 50 μm into four equal sets, collected in 6-well plates filled with ice-cold PBS and washed with PBS three times at room temperature. The following incubation steps were performed with mild shaking on an orbital shaker, unless stated otherwise. Brain sections were incubated in 1% Triton X-100 in PBS at room temperature for 1 h. The sections were blocked with 10% Blocking One (Nacalai Tesque) in 0.3% Triton X-100-treated PBS (blocking solution) for 1 h at room temperature without shaking. The sections were incubated in primary antibodies diluted with blocking solution (dilutions and types of each antibody are listed below) at 4 °C overnight, then washed three times, incubated with secondary antibodies at 4 °C overnight, washed with PBS, then mounted and coverslipped with HardSet Antifade Mounting Medium with DAPI (Vectashield).

The primary antibodies used in this study were; rabbit anti-cFOS (1:4,000, ABE457, Millipore), goat anti-mCherry (1:15,000, AB0040-200, Sicgen), rat anti-GFP (1:5,000, 04404-84, Nacalai Tesque), mouse anti-TH (1:1,000, sc-25269, Santa Cruz Biotechnology), mouse anti-orexin-A (1:200, sc-80263, Santa Cruz Biotechnology) and rabbit anti-MCH (1:2,000, M8440, Sigma). The secondary antibodies were: Alexa Fluor 488 donkey anti-rat, 488 donkey anti-rabbit, 594 donkey anti-rabbit, 594 donkey anti-goat, 647 donkey anti-mouse and 647 donkey anti-rabbit (1:1,000, Invitrogen). For Nissl staining, sections were counterstained with NeuroTrace 435/455 blue fluorescent Nissl stain (1:500, N-21479, Invitrogen) during the secondary antibody step, and coverslipped with FluorSave Reagent (Millipore). Brain sections were observed using an Axio Zoom.V16 (Zeiss) and a TCS SP8 laser confocal microscope (Leica). Brain regions were defined according to the brain maps adapted from Paxinos and Franklin's atlas of the mouse brain[38] and Paxinos and Watson's atlas of the rat brain[39]. Some maps in these books were modified to simply depict stereotaxic orientation.

In situ hybridization

Fluorescence in situ hybridization was performed with the RNAscope Fluorescent Multiplex Kit (Advanced Cell Diagnostics) using the following probes: Probe-Mm-*Qrfp* 464341, *mCherry* 431201-C2, Mm-*Slc32a1* 319191-C3, Mm-*Slc17a6* 319171, Mm-*Adcyap1* 405911-C2, Mm-*Bdnf*-CDS 457761-C3, Mm-*Ptger3*-O1 501831-C3 and Rn-*Qrfp* 834441.

Mice or rats were perfused and fixed in fresh 4% PFA and brains were processed until sectioning by the same method as described in 'Immunohistochemistry' except using DEPC-treated PBS. Brains were sectioned coronally into six sections (20 μm) with a cryostat (Leica), and mounted on Superfrost Plus microscope slides (Fisherbrand). Pretreatment procedures (post-fixation and dehydration) and RNAscope fluorescent multiplex assay were performed following the RNAscope Assay Guide (document numbers 320513 and 320293). For the combination with immunohistochemistry, after the amplification phase, samples were immediately moved into the washing phase of immunohistochemistry using slide vats. Then, the slides were processed with the same procedure as was used for immunohistochemistry.

Cell counting

Images of FOS experiments (Fig. 1c) and fluorescence in situ hybridization (Fig. 4a, Extended Data Figs. 11d, 12f) were obtained with a TCS SP8 laser confocal microscope (Leica), and *z*-stacked to 25 μm and 10 μm, respectively. Images were counted using image-analysing software (LAS X, Leica). In Fig. 1c, mCherry-expressing Q neurons and FOS signals were quantified by counting 6 sections from mice ($n = 3$) per group (0.50 and 0.26 mm from bregma). In Fig. 4a, Q neurons, *Vglut2*- and *Vgat*-expressing cells were quantified by counting 6 sequential sections from Q-hM3D mice ($n = 2$ slices of 0.50, 0.38 and 0.26 mm from bregma as shown in the figure). In Extended Data Fig. 11d, Q neurons, *Adcyap1*-, *Bdnf*- and *Ptger3*-expressing cells were calculated by counting 6 sections from Q-hM3D mice ($n = 2$; same series of slices as Fig. 4a).

Retrograde tracing

Male $Qrfp^{iCre}$ mice (10–12 weeks old) were injected with viruses as shown below. First, 0.16 μl AAV_{10}-DIO-TVA-mCherry and 0.33 μl AAV_{10}-DIO-RG were delivered to express TVA–mCherry and rabies glycoprotein (RG) in Q neurons in the AVPe/MPA (AP, 0.38 mm; ML, 0.25 mm; DV, −5.25 mm from bregma) unilaterally. Three weeks later, 0.3 μl SADΔG-GFP(EnvA) was injected at the same site. Six days later, mice were fixed and treated according to the immunohistochemistry procedure. Whole brain sections were observed to detect starter (mCherry and GFP double-positive) neurons and input (GFP-positive) neurons using an Axio Zoom.V16 (Zeiss) and a TCS SP8 laser confocal microscope (Leica).

Optogenetic manipulation

Mice were connected with optical fibre patch cable (200-µm diameter; NA: 0.22, 1.0 m long; Doric Lenses). We used DPSS lasers (473-nm blue or 589-nm yellow; Shanghai Laser) to apply optogenetic manipulations. The laser power at the optic fibre tip was adjusted to 8–10 mW. In Fig. 2c–e, 473-nm laser stimulation was applied at 1 Hz for 1 s every 30 min (repeated 4 times), controlled by a TTL pulse generator (Amuza). In Extended Data Fig. 7e, a 473-nm laser was applied at 1 Hz for 1 s every 60 min (repeated 3 times) and a 589-nm laser was applied at 1 Hz for 5 s 3 min after the second stimulation with the 473-nm laser. Mice with incorrect fibre placement were excluded from data analysis.

Behavioural tests

All behavioural tests were performed during the dark phase, with five Q-hM3D mice and five *Qrfp*^{iCre} mice that were injected with AAV$_{10}$-DIO-mCherry into the AVPe/MPA. After AAV injection, mice were housed singly in home cages and allowed to recover for two weeks, then acclimated to the experimenter's handling. All tests except the rotarod test were recorded with a visual video camera (FDR-AX60, Sony) and analysed by Smart Video Tracking Software (Panlab, Harvard Apparatus). We performed behavioural tests in the following order, one test each day.

Open-field tests were performed by using a square open-field arena (made of opaque plastic, W40 × D40 × H30 cm) with dim light (less than10 lx). Each mouse was placed in and allowed to freely explore the arena for 20 min. The arena was wiped with 70% ethanol and weakly acidic water after each session.

Novel object recognition tests were carried out in the same arena as the open-field tests. The mice were put in the centre of the arena, allowed to explore for 20 min and to touch two identical objects (object A) placed symmetrically. One day later, the mice were put back in the arena for the novel object recognition test trial. For the test trials, one of the previous familiar objects (object A) remained in the arena, but the other one was replaced with a novel object (object B). The time spent in the area of each object (a 5-cm-diameter circle with the object at the centre) was measured for calculating a discrimination ratio defined as follows: (Time B − Time A)/(Time A + Time B). The objects were a 25-ml cell culture flask filled with sand; and stacked plastic blocks. Flasks and blocks were randomly assigned as object A or B in each experiment.

Elevated plus maze tests were conducted on an apparatus made of white plastic, consisting of a central area (5 × 5 cm), two open arms (25 × 5 cm) and two closed arms (25 × 5 cm) with 25-cm high walls with 10 lx illumination. The mouse was put in the central area and allowed to explore for 15 min.

Rotarod tests were performed using an accelerating rotarod (Ugo Basile) in which a mouse was placed on the rotating drum (3-cm diameter). The initial speed of the rotarod was set at 4 rpm. The speed gradually increased from 4 to 40 rpm over 300 s.

For tail suspension tests, mice were suspended from their tails by a strip of masking tape that was placed approximately 2 cm from the tip of the tail for 5 min.

A week after behavioural tests, all mice were fixed, and haematoxylin and eosin staining was performed using the standard method.

Measuring daily body weight, food intake and locomotor activity

We measured daily body weight, food intake and locomotor activity before and after QIH induction in six Q-hM3D mice. Body weight and amount of chow consumed were measured at the beginning of the dark period. On days 4 and 6, saline and CNO (1 mg kg^{-1}) were administrated intraperitoneally at a volume of 10 µl g^{-1}, respectively. Locomotor activity was constantly detected throughout all experimental days with a customized device for sensing object locomotion, which was placed 25 cm above the home cage and enclosed in a sound attenuating chamber (Muromachi). The data were analysed with SOF-860 software (Med Associates) and retrieved every hour.

Electrophysiological analysis

Mice were decapitated under deep anaesthesia with isoflurane (Pfizer). Brains were extracted and cooled in ice-cold cutting solution containing the following: 125 mM choline chloride, 25 mM $NaHCO_3$, 10 mM D(+)-glucose, 7 mM $MgCl_2$, 2.5 mM KCl, 1.25 mM NaH_2PO_4 and 0.5 mM $CaCl_2$ bubbled with O_2 (95%) and CO_2 (5%). Horizontal brain slices (250-µm thickness), including the hypothalamus, were prepared with a vibratome (VT1200S, Leica) and maintained for 1 h at room temperature in artificial cerebrospinal fluid (ACSF) containing the following: 125 mM NaCl, 26 mM $NaHCO_3$, 10 mM D(+)-glucose, 2.5 mM KCl, 2 mM $CaCl_2$ and 1 mM $MgSO_4$ bubbled with O_2 (95%) and CO_2 (5%). The electrodes (5–8 MΩ) were filled with an internal solution containing the following: 125 mM K-gluconate, 10 mM HEPES, 10 mM phosphocreatine, 0.05 mM tolbutamide, 4 mM NaCl, 4 mM ATP, 2 mM $MgCl_2$, 0.4 mM GTP and 0.2 mM EGTA, pH 7.3, adjusted with KOH). Firing of hM3Dq-mCherry-expressing neurons was recorded in the current-clamp mode at a temperature of 30 °C. CNO (1 µM) was bath-applied to examine the effects. The combination of a MultiClamp 700B amplifier, Digidata 1440A A/D converter and Clampex 10.3 software (Molecular Devices) was used to control membrane voltage and data acquisition.

Induction of FIT

Each FIT-induction experiment was designed to record the metabolism of the mouse for at least three days. The mice were introduced to the chamber the day before recording started (day 0). Food and water were freely available. The ambient temperature was set as indicated on day 0 and kept constant throughout the experiment. A telemetry temperature sensor implanted in the mouse was turned on before placing it in the chamber. The standard experimental design was as follows: on day 2, zeitgeber time (ZT) 0, food was removed to induce torpor. After 24 h, on day 3, at ZT0, food was returned to each animal.

Three-dimensional imaging of transparent mouse brain

Transparent mouse brains were generated by the Sca*l*eS method as described previously[40]. ScaleS solutions were made using urea crystals (Wako Pure Chemical Industries, 217-00615), D(−)-sorbitol (Wako Pure Chemical Industries, 199-14731), methyl-β-cyclodextrin (Tokyo Chemical Industry, M1356), γ-cyclodextrin (Wako Pure Chemical Industries, 037-10643), *N*-acetyl-L-hydroxyproline (Skin Essential Actives, Taiwan), dimethyl sulfoxide (DMSO) (Wako Pure Chemical Industries, 043-07216), glycerol (Sigma, G9012) and Triton X-100 (Nacalai Tesque, 35501-15). Brains of *Qrfp*^{iCre} mice injected with AAV$_9$-hSYN-DIO-GCaMP6 s were fixed and cleared with ScaleS. Images were obtained with a laser confocal microscope (Olympus, FV1200 with XLSLPN25XGMP (NA 1.00, WD: 8 mm) (RI: 1.41–1.52)).

Statistical analysis

In this study, we used Bayesian statistics to evaluate our hypothesis and experimental results. We designed a statistical model with parameters representing the structure of the hypothesis and fitted the model to the experimental results. Bayesian inference estimates the posterior probability distribution of the model parameters from the likelihood distribution and prior probability distribution of the parameters. The posterior distributions provide information on how the model can describe the hypothesis from the experimental results. Bayesian modelling can explicitly include all types of uncertainty; therefore, it can deal with data with noise in the observation or it can fully use information from a small number of samples that potentially have a wide range of uncertainty. Furthermore, it can deal with multiple layers of multiple groups with different numbers of samples using hierarchical models.

All of these advantages of Bayesian inference render it an appropriate method for handling commonly seen issues in animal experiments. Model fitting was performed using Hamiltonian Monte Carlo with its adaptive variant, the No-U-turn Sampler, as implemented in Stan v.2.18.0 with the RStan library[41] in R v.3.52[42]. We assessed convergence by inspection of the trace plots, Gelman and Rubin's convergence diagnostic and an estimate of the effective number of samples. The model priors were defined to be weakly informative and conservative, which are specified in the following sections. The fundamental principles and techniques for designing the statistical models were based on a previous publication[43]. The source code for the models and the data used for analysis are available at https://briefcase.riken.jp/public/JjtgwAnqQ8lAgyI.

Body weights of $Qrfp^{iCre}$ mice were modelled at a given age and strain by a state-space hierarchical model (Extended Data Fig. 1d, e, code folder QRFP_KO_BW). Mice in each group (wild-type mice ($n = 9$), heterozygous ($n = 9$) and homozygous ($n = 10$) $Qrfp^{iCre}$ mice) were raised in each cage without identification of individuals. When the unobservable baseline of body weight is defined as a time-variable $B_{t,s}$, in which t is the time point and s is the index of strains (1, 2 and 3 for wild-type, heterozygous and homozygous $Qrfp^{iCre}$ mice, respectively), with the trend $\eta_{t,s}$ and the total time point T, the observed state $Y_{t,i}$ can be described by modelling the observational error by a log-normal distribution as:

$$Y_{t,s} \approx \text{log-normal}\left[\log(B_{t,s}) - \frac{(\sigma_1)^2}{2},\ \sigma_1\right] \tag{1}$$

$$\begin{cases} B_{1,s} = +\eta_1,\ t = 1 \\ B_{2,s} - B_{1,s} = B_{1,s} + \eta_{2,s},\ t = 2 \\ B_{t,s} - B_{t-1,s} = B_{t-1,s} - B_{t-2,s} + \eta_{t,s},\ t \geq 3 \end{cases} \tag{2}$$

$$\eta_{t,s} \approx \text{Normal}(0, \sigma_2) \tag{3}$$

$$t = 1 \cdots T \tag{4}$$

$$s = \{1, 2, 3\} \tag{5}$$

Uniform priors were applied for every parameter except σ_1 and σ_2, which were drawn from standard half-normal distribution.

The spiking frequency of *Qrfp*-positive neurons in brain slices was modelled by parameterizing the difference in spiking frequency when neurons were activated by CNO (Extended Data Fig. 3c–e, code folder Patch_M3_CNO). When the total number of slices is K, and the observed spiking frequencies of the control and CNO-administered recording of the ith slice are B_i and C_i, respectively, B_i is modelled by β_{BASE} with observational errors, and C_i is modelled by the sum of β_{BASE} and β_{CNO} with observational errors. Because spiking frequency is a positive real number, errors can be modelled by a log-normal distribution; therefore, B_i and C_i can be described as:

$$B_i \approx \text{log-normal}\left(\log(\beta_{\text{BASE}}) - \frac{(\sigma_{\text{ERROR}})^2}{2},\ \sigma_{\text{ERROR}}\right) \tag{6}$$

$$C_i \approx \text{log-normal}\left(\log(\beta_{\text{BASE}} + \beta_{\text{CNO}}) - \frac{(\sigma_{\text{ERROR}})^2}{2},\ \sigma_{\text{ERROR}}\right) \tag{7}$$

$$\beta_{\text{BASE}} \approx \text{Normal}(0, \sigma_{\text{BASE}}) \tag{8}$$

$$\beta_{\text{CNO}} \approx \text{Normal}(0, \sigma_{\text{CNO}}) \tag{9}$$

$$i = 1 \cdots K \tag{10}$$

All σ values were sampled from standard half-normal distributions.

T_{BAT} values of optogenetically or chemogenetically stimulated mice were modelled in a hierarchical multilevel model (Figs. 2d, 4c; code folders TBAT_Opt and TBAT_Flox). Four groups of mice were included in this experiment. T_{BAT} was recorded at 1 Hz, and every 10 s the 10-s maximum was stored for further analysis. Ten-minute recording of every T_{BAT} of interest was included in the analysis. When K is the total number of mice, and Y is T_{BAT} during the duration of interest of mouse j that belongs to group i, Y can be described as the sum of the global mean parameter β, the group parameter β_{GROUP} and the individual mouse parameter β_{MOUSE} with the observational noise modelled in a Cauchy distribution of a scale parameter σ_{ERROR} as:\

$$Y_{i,j} \approx \text{Cauchy}\Big(\beta + \beta_{\text{GROUP}[i]} + \beta_{\text{MOUSE}[j]},\ \sigma_{\text{ERROR}}\Big) \tag{11}$$

$$\beta_{\text{GROUP}} \approx \text{Normal}(0,\ \sigma_{\text{GROUP}}) \tag{12}$$

$$\beta_{\text{MOUSE}} \approx \text{Normal}(0,\ \sigma_{\text{MOUSE}}) \tag{13}$$

$$i = \{1,\ 2,\ 3,\ 4\} \tag{14}$$

$$j = 1 \cdots K \tag{15}$$

All σ values were sampled from standard half-normal distributions. Differences in T_{BAT} among groups were compared by estimating the mean T_{BAT} for each group from posterior distributions, which is the sum of β and β_{GROUP} with normally distributed noise at a standard deviation of σ_{MOUSE}.

To evaluate the thermoregulatory system during QIH and normal conditions, heat loss and production of the animal was described in a hierarchical multilevel model (Fig. 3c–k, code folder QIH_GTRH). Three parameters (G, T_{R} and H) during two metabolic conditions (normal and QIH) were estimated from the metabolically stable state of the animal at various ambient temperatures. The detailed methods have previously been described[5]. In brief, a linear model consisting of the controllable parameter T_{A} and the observable parameters T_{B} and VO_2 was fitted to the experimental results for both T_{B} and VO_2 using T_{A} as a predictor with normally distributed noise. The posterior distribution of the slope and intercept coefficients for each model were then used to estimate G, T_{R} and H. For estimation of T_{R} and H, the model is designed for monotonically increasing T_{B} and decreasing VO_2 against T_{A}. Therefore, the parameter estimation during QIH used T_{B} and VO_2 during $T_{\text{A}} = 16$ °C, 20 °C and 24 °C (Fig. 3b, Extended Data Fig. 8b). In this analysis, priors of the standard deviation of the noise were standard half-normal distributions, and the other parameters used the positive region of uniform distribution except the intercept coefficient of T_{B}, which used uniform distribution owing to possible negative values.

The circadian transition of metabolism in Q-TeTxLC mice was analysed by modelling the metabolism by clustering the recorded values into the light phase (L-phase) and the dark phase (D-phase) (Extended Data Fig. 10h, code folder TeTxLC_LD). Specifically, when Y is the observed T_{B} of group i at phase j, Y can be described as the sum of base metabolism (light-phase metabolism) and the difference between the dark phase with normally distributed observational noise as:

$$Y_{i,\ j} \approx \text{Normal}\Big(\beta_{\text{BASE}[i]} + \beta_{\text{DARK}[i]} P_j,\ \sigma_{\text{ERROR}}\Big) \tag{16}$$

$$\beta_{\text{BASE}} \approx \text{Normal}(0,\ \sigma_{\text{BASE}}) \tag{17}$$

$$\beta_{\mathrm{DARK}} \approx \mathrm{Normal}(0, \sigma_{\mathrm{DARK}}) \tag{18}$$

$$i = \{1\text{: control},\ 2\text{: TeTxLC}\} \tag{19}$$

$$j = \{1\text{: L-phase},\ 2\text{: D-phase}\} \tag{20}$$

$$\begin{cases} P_1 = 0 \\ P_2 = 1 \end{cases} \tag{21}$$

All σ values were sampled from standard half-normal distributions. For modelling VO_2, the fundamental model structure was identical to T_B modelling except that the observational error was modelled as a log-normal distribution because VO_2 assumes only positive real numbers.

Statistics and reproducibility

The number of animals or samples used in each experiment are stated in the manuscript or in the figures. The numbers of experimental repetitions were as follows: Fig. 1a, 4 times (coronal) and 3 times (horizontal); Fig. 1b, 12 times; Fig. 1c, 3 times; Fig. 2a, b, 3 times; Fig. 2e, 6 times; Fig. 3a: numbers of mice are: 6, 11, 5, 5, 6, 6 and 6 for T_A = 8, 12, 16, 20, 24, 28 and 32 °C, respectively; Fig. 3l, 4 times; Fig. 4a, 3 times; Extended Data Fig. 1b, c, 3 times; Extended Data Fig. 2b, twice (horizontal) and 3 times (coronal); Extended Data Fig. 3, twice (a) and 9 times (b); Extended Data Fig. 4, twice; Extended Data Fig. 5b, 5 times; Extended Data Fig. 6e, twice; Extended Data Fig. 6f, 4 times; Extended Data Fig. 7b–d, 3 times; Extended Data Fig. 10, once (b) and 3 times (c); Extended Data Fig. 11, 3 times (b, c) and twice (d, e); Extended Data Fig. 12, 4 times (c) and twice (e, f); Supplementary Video 2, twice; Supplementary Video 3, 6 times; and Supplementary Video 4, 4 times.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Data used for Bayesian estimation are included with the source code. Other data are available from the corresponding authors on request.

Code availability

All of the source code of the models used for Bayesian estimation is available at https://briefcase.riken.jp/public/JjtgwAnqQ8lAgyI. See 'Statistical analysis' in Methods for details.

33. Mieda, M. et al. Cellular clocks in AVP neurons of the SCN are critical for interneuronal coupling regulating circadian behavior rhythm. *Neuron* **85**, 1103–1116 (2015).
34. Osakada, F. & Callaway, E. M. Design and generation of recombinant rabies virus vectors. *Nat. Protoc.* **8**, 1583–1601 (2013).
35. Saito, Y. C. et al. Monoamines inhibit GABAergic neurons in ventrolateral preoptic area that make direct synaptic connections to hypothalamic arousal neurons. *J. Neurosci.* **38**, 6366–6378 (2018).
36. Gehrmann, J. et al. Phenotypic screening for heart rate variability in the mouse. *Am. J. Physiol. Heart Circ. Physiol.* **279**, H733–H740 (2000).
37. Sunagawa, G. A. et al. Mammalian reverse genetics without crossing reveals *Nr3a* as a short-sleeper gene. *Cell Rep.* **14**, 662–677 (2016).
38. Franklin, K. B. J. & Paxinos, G. *The Mouse Brain in Stereotaxic Coordinates* 3rd edn (Academic Press, 2007).
39. Paxinos, G. & Watson, C. *The Rat Brain in Stereotaxic Coordinates* 7th edn (Academic Press, 2013).
40. Hama, H. et al. Scal*e*S: an optical clearing palette for biological imaging. *Nat. Neurosci.* **18**, 1518–1529 (2015).
41. Stan Development Team. RStan: the R interface to Stan. R package v.2.19.3 (2020).
42. R Core Team. R: A language and environment for statistical computing, v.2.18.0 (R Foundation for Statistical Computing, 2018).
43. McElreath, R. *Statistical Rethinking: a Bayesian Course with Examples in R and Stan* 1st edn (CRC Press, 2016).
44. Romanov, R. A. et al. Molecular interrogation of hypothalamic organization reveals distinct dopamine neuronal subtypes. *Nat. Neurosci.* **20**, 176–188 (2017).
45. Wickersham, I. R. et al. Monosynaptic restriction of transsynaptic tracing from single, genetically targeted neurons. *Neuron* **53**, 639–647 (2007).
46. Wang, T. A. et al. Thermoregulation via temperature-dependent PGD_2 production in mouse preoptic area. *Neuron* **103**, 309–322 (2019).
47. Tan, C. L. & Knight, Z. A. Regulation of body temperature by the nervous system. *Neuron* **98**, 31–48 (2018).
48. Lazarus, M. et al. EP3 prostaglandin receptors in the median preoptic nucleus are critical for fever responses. *Nat. Neurosci.* **10**, 1131–1133 (2007).

Acknowledgements This study was supported by a JSPS KAKENHI Grant-in-Aid for Scientific Research (B) (JP 18H02595) (T.S.); a JSPS KAKENHI Grant-in-Aid for Scientific Research on Innovative Areas 'Willdynamics' (16H06401); JSPS KAKENHI grant number JP19K22465 (T.S.); JST CREST grant number JPMJCR1655 Japan (T.S.); the RIKEN Special Postdoctoral Researcher program (G.A.S.); a JSPS KAKENHI Grant-in-Aid for Scientific Research on Innovative Areas 'Thermal Biology' (18H04706) (G.A.S.); a JSPS KAKENHI Grant-in-Aid for Scientific Research (A) (19H01066) (G.A.S.); a research grant from the Astellas Foundation for Research on Metabolic Disorders (G.A.S.); and JSPS KAKENHI grant number 19J20876 (T.M.T.). We thank the animal resource centre at the University of Tsukuba and LARGE, RIKEN BDR for housing the mice; Y. Cherasse for preparing virus vectors; A. Miyasaka and Y. Niwa for discussion; and W. Gray for proofreading the manuscript. All of the brain diagrams used in figures were made based on the illustrations in Paxinos and Franklin's atlas of the mouse brain[38] and Paxinos and Watson's atlas of the rat brain[39].

Author contributions T.M.T., G.A.S. and T.S. conceived the project and designed the experiments. T.M.T. performed all stereotaxic surgeries and tissue samplings and conducted experiments except as noted. T.M.T., K.I., M.T. and G.A.S. performed metabolism and biological signal-recording experiments. T.M.T. and S.S. performed histology analyses. M.A., M.Y., T.S. and K. Sakimura designed and generated genetically modified mice. T.S. and K. Sakurai. prepared AAV vectors. T.S. prepared rabies virus vectors. H.H. and A.M. performed three-dimensional imaging of transparent brains. E.H. performed electrophysiology experiments. G.A.S. performed the statistical analyses. T.M.T., G.A.S. and T.S. wrote the manuscript, with input from all co-authors.

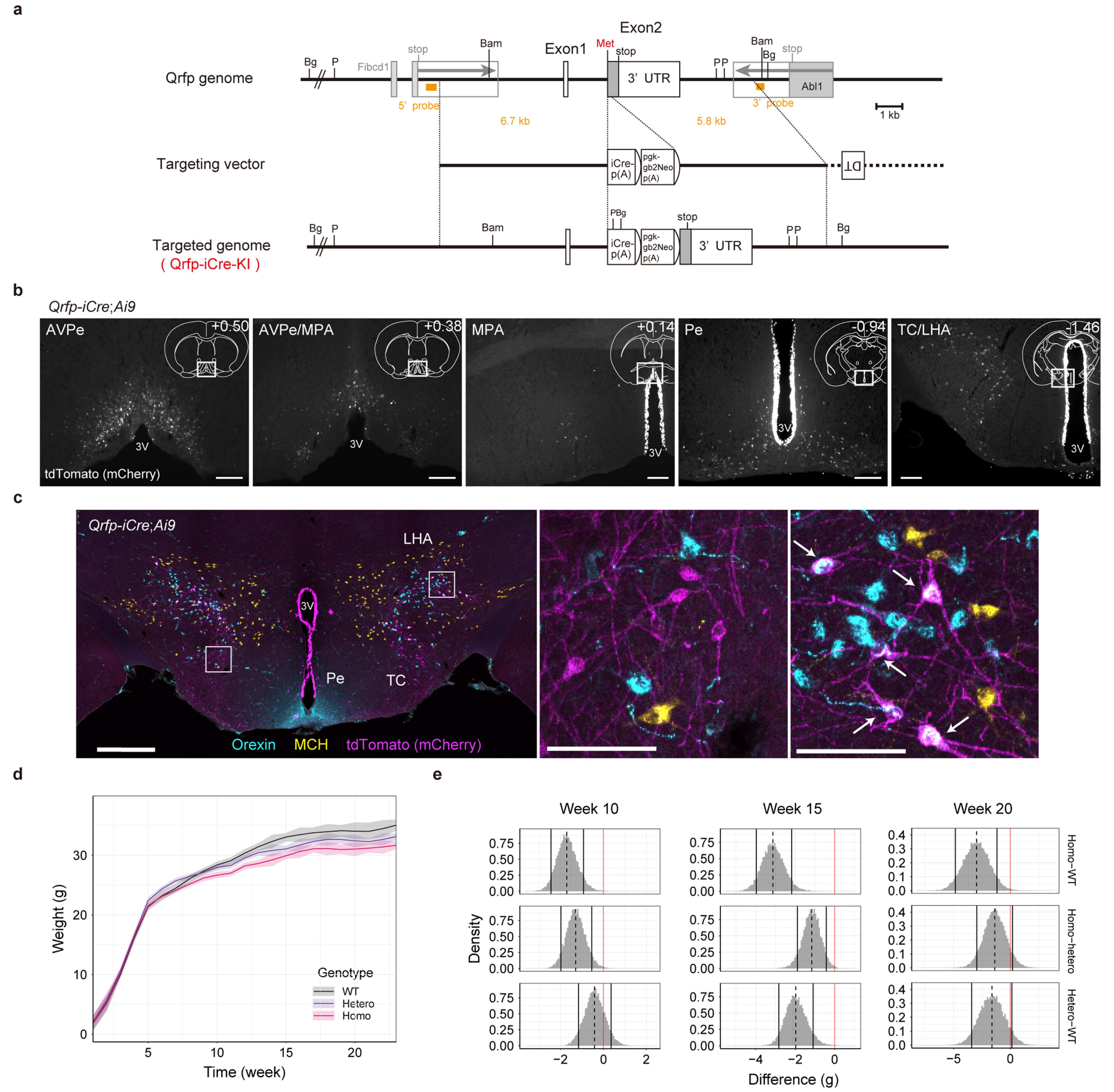
Competing interests The authors declare no competing interests.

Additional information

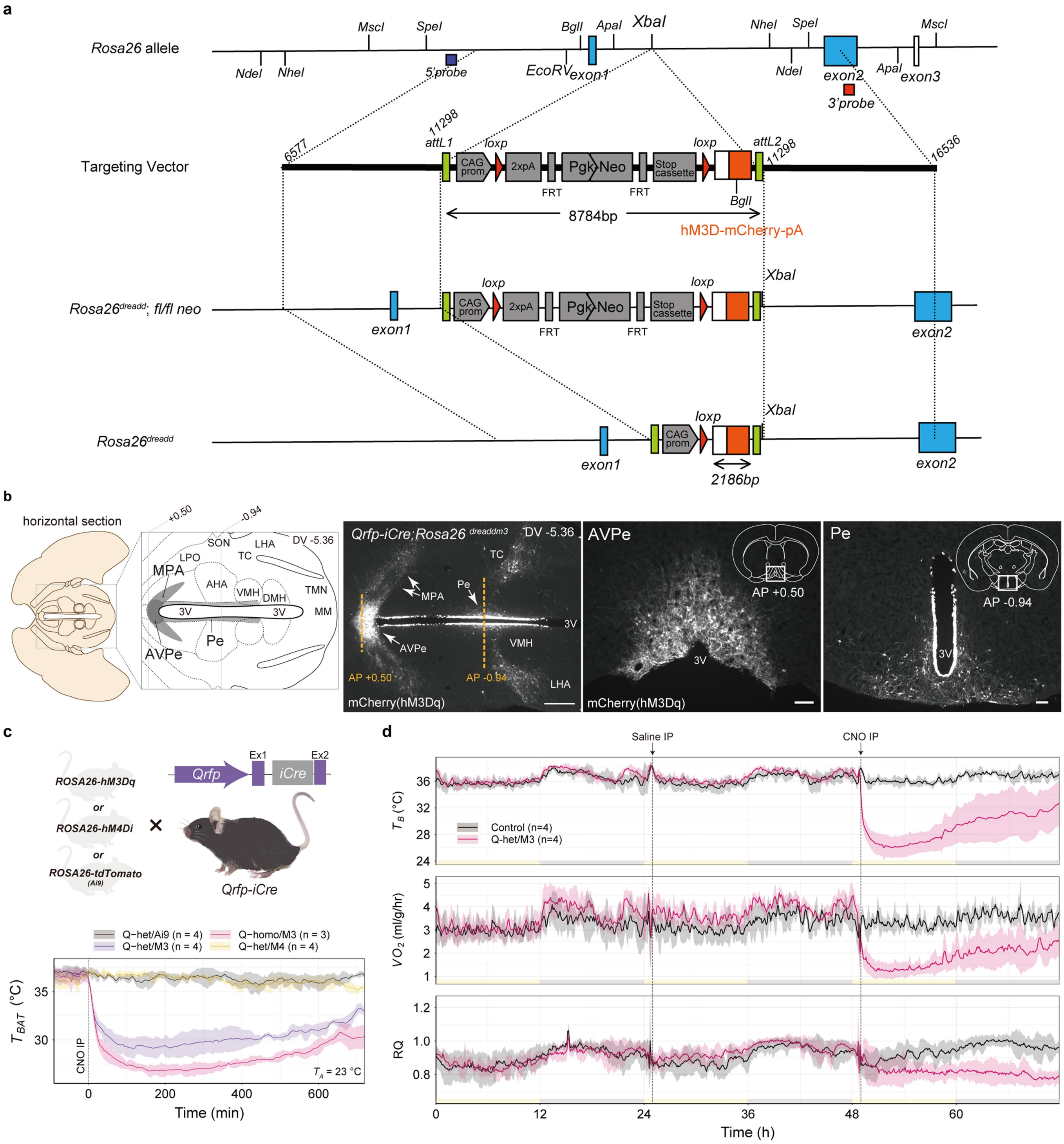
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2163-6.

Correspondence and requests for materials should be addressed to G.A.S. or T.S.

Peer review information *Nature* thanks Rob Henning, Shaun Morrison, Richard Palmiter and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

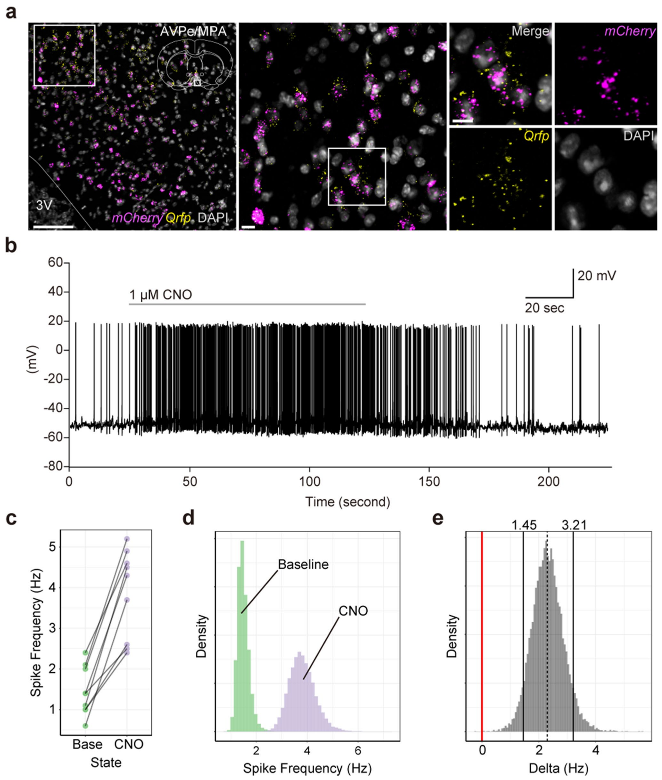


Extended Data Fig. 1 | Generation of *Qrfp^{iCre}* mice. To examine the role of QRFP-producing neurons, we engineered mice in which codon-improved Cre recombinase (*iCre*) is inserted in the *Qrfp* allele. **a**, Targeting vector and structure of the targeted allele of *Qrfp^{iCre}* mice. We mated mice with the targeted genome with FLP66 mice to delete the pgk-Neo cassette and create the *Qrfp^{iCre}* mice used in this study. **b**, Distribution of Cre-positive neurons in coronal sections of brain prepared from *Qrfp^{iCre};Ai9* mice. Scale bars, 200 µm. **c**, Immunostaining of hypothalamic slices prepared from a *Qrfp^{iCre};Ai9* mouse with anti-mCherry, anti-orexin and anti-melanin-concentrating hormone (MCH) antibodies. Along the wall of the third ventricle, we found extensive expression of mCherry, presumably derived from tanycytes and ependymal cells. However, we could not express exogenous genes by injecting Cre-dependent AAV vectors into this region in adult mice, suggesting transient expression of Cre in these cells during the developmental stage. In addition, we observed the existence of iCre-positive neurons in the LHA in reporter mice crossed with *Qrfp^{iCre}* mice that were also positive for orexin-like immunoreactivity—although a previous study did not find orexin and QRFP double-positive cells in adult mice[13]. This suggests that a low level of iCre is expressed in some orexin neurons, and that orexin neurons and QRFP neurons might be derived from the same cell lineage. Single-cell RNA-sequencing analysis of the hypothalamus showed colocalization of *Qrfp* and *Orexin* (also known as *Hcrt*), and hierarchical clustering defined by molecular fingerprints showed that orexin- and QRFP-expressing neurons have a close neuronal lineage[44]. The middle and right images are magnifications of the boxed areas. QRFP-expressing neurons in the LHA were positive for mCherry (arrows) but negative for MCH. Scale bars, 500 µm (left); 100 µm (middle, right). **d**, Growth curve of *Qrfp^{iCre}* mice ($n = 9$ wild type (WT), $n = 9$ *Qrfp^{iCre}* heterozygous and $n = 10$ *Qrfp^{iCre}* homozygous). Lines show median and shaded areas denote the estimated 89% HPDI of the body weight of each group at a given age. **e**, Posterior distribution of estimated difference in body weight between two groups. The dotted line shows median and solid lines denote 89% HPDI of differences. Homozygous *Qrfp^{iCre}* mice are smaller than wild-type mice, consistent with a previous observation[13].

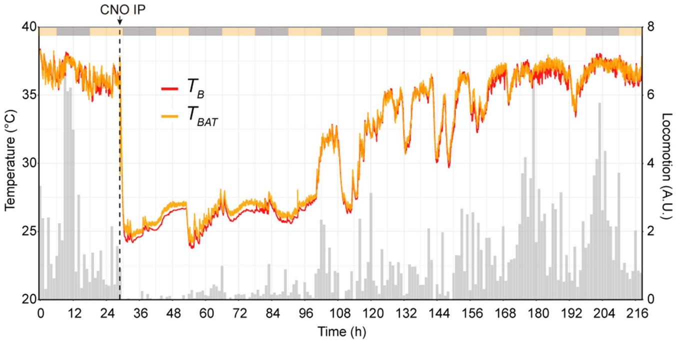


Extended Data Fig. 2 | See next page for caption.

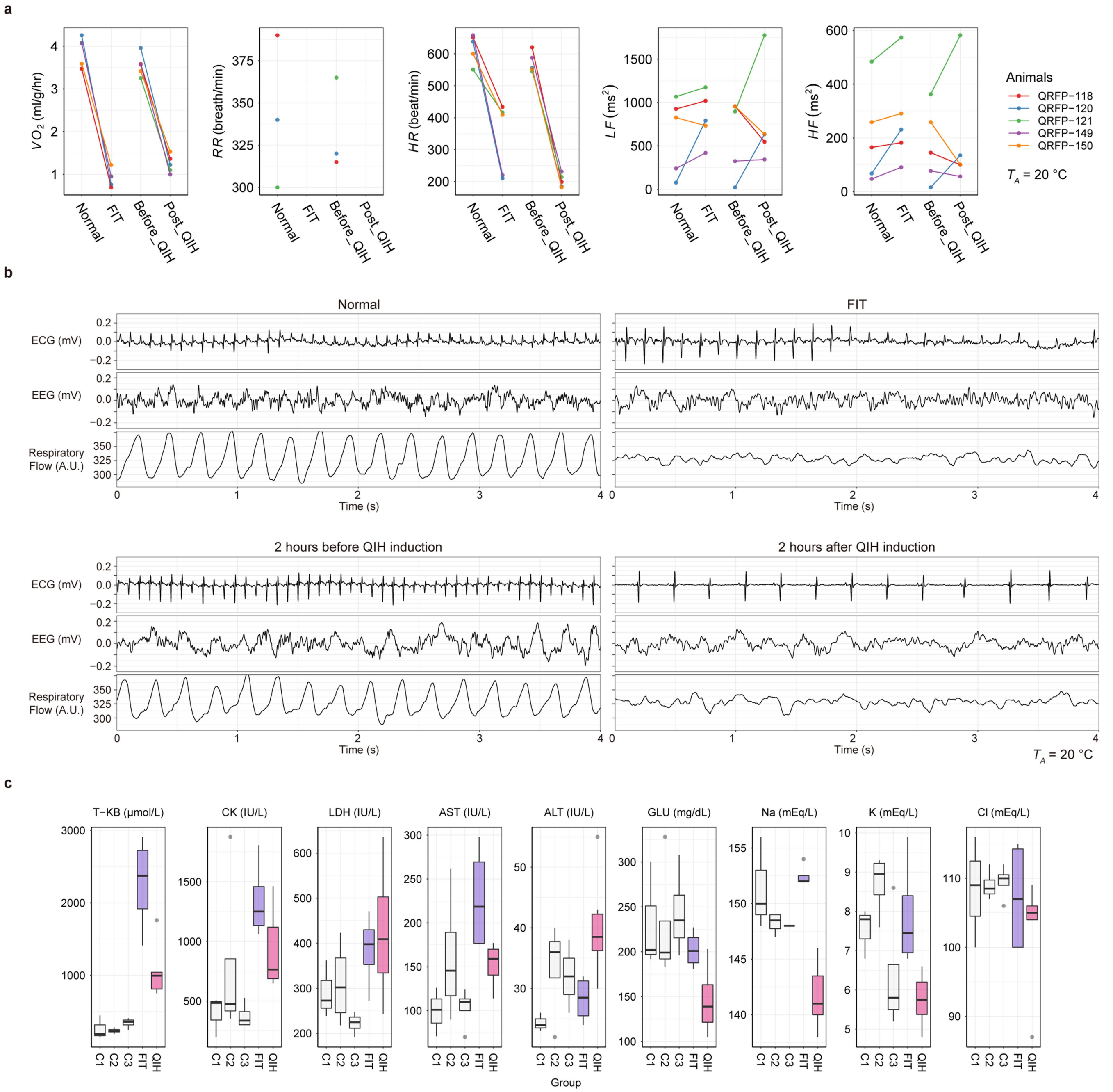
Extended Data Fig. 2 | Expression of DREADD receptors in *Qrfp*^{iCre} neurons. We generated *Rosa26*^{dreaddm3} mice and crossed them with *Qrfp*^{iCre} mice (Extended Data Fig. 1) to obtain mice that express hM3Dq-mCherry exclusively in iCre-expressing cells (*Qrfp*^{iCre};*Rosa26*^{dreaddm3} mice). **a**, Generation of mice that express hM3Dq and hM4Di in Cre-expressing neurons. Targeting vectors and structures of the targeted alleles of *Rosa26*^{dreaddm3} and *Rosa26*^{dreaddm4} mice. We mated these mice with FLP66 mice to delete the pgk-Neo cassette. Orange boxes indicate hM3Dq-mCherry or hM4Di-mCherry. Because the CAG promoter drives expression of hM3Dq-mCherry or hM4Dq-mCherry only after Cre-mediated excision of the floxed stopper element, this allowed us to express hM3Dq or hM4Di specifically in Cre-expressing neurons. **b**, Horizontal and coronal sections of brain prepared from a *Qrfp*^{iCre};*Rosa26*^{dreaddm3} mouse, showing the distribution of mCherry-positive neurons in the hypothalamus. **c**, Top, strategy for chemogenetic excitation or inhibition of whole iCre-positive neuronal populations in *Qrfp*^{iCre} mice. Bottom, chemogenetic excitation of iCre-positive cells in *Qrfp*^{iCre} mice induced hypothermia. Heterozygous (Q-het) or homozygous (Q-homo) *Qrfp*^{iCre} mice with heterozygous *Rosa26*^{dreaddm3} (M3) and/or *Rosa26*^{dreaddm4} (M4) alleles were subjected to experiments. CNO was administered at ZT12 (start of the dark period). The ambient temperature was 23 °C. We found that excitatory manipulation of *Qrfp*^{iCre} neurons in mice resulted in severe immobility. As the posture of these mice was similar to that observed during daily torpor, we initially postulated that activation of iCre-positive cells induced a daily torpor-like state. To evaluate this hypothesis, we measured body temperature and found that the induced state of immobility was accompanied by marked, long-lasting hypothermia. T_{BAT} decreased beginning about 5 min after CNO administration and lasted 12 h. Mice spontaneously recovered without external warming. By contrast, inhibitory DREADD manipulation of iCre-positive neurons did not have any effect on T_{BAT}. Notably, hM3Dq-mediated activation of iCre-positive neurons in *Qrfp*^{iCre};*Rosa26*^{dreaddm3} mice induced robust hypothermia, even in homozygous *Qrfp*^{iCre} mice in which *Qrfp* sequences are completely replaced by *iCre* in both alleles. This suggests that QRFP itself does not have a role in inducing hypothermia. The degree of hypothermia was greater in QRFP-deficient mice, which indicates that endogenous QRFP itself counteracts the hypothermia. **d**, Excitatory manipulation of Q neurons in *Qrfp*^{iCre};*Rosa26*^{dreaddm3} mice in the light period (at ZT1) also induced a long-lasting hypothermic state ($n = 4$ mice for each condition). Line and shading in **c**, **d** denote mean and s.d. of each group. AHA, anterior hypothalamus; ARC, arcuate nucleus; LPO, lateral preoptic area; MM, medial mammillary nucleus; SON, supraoptic nucleus; TMN, tuberomammillary nucleus; VMH, ventromedial hypothalamus.



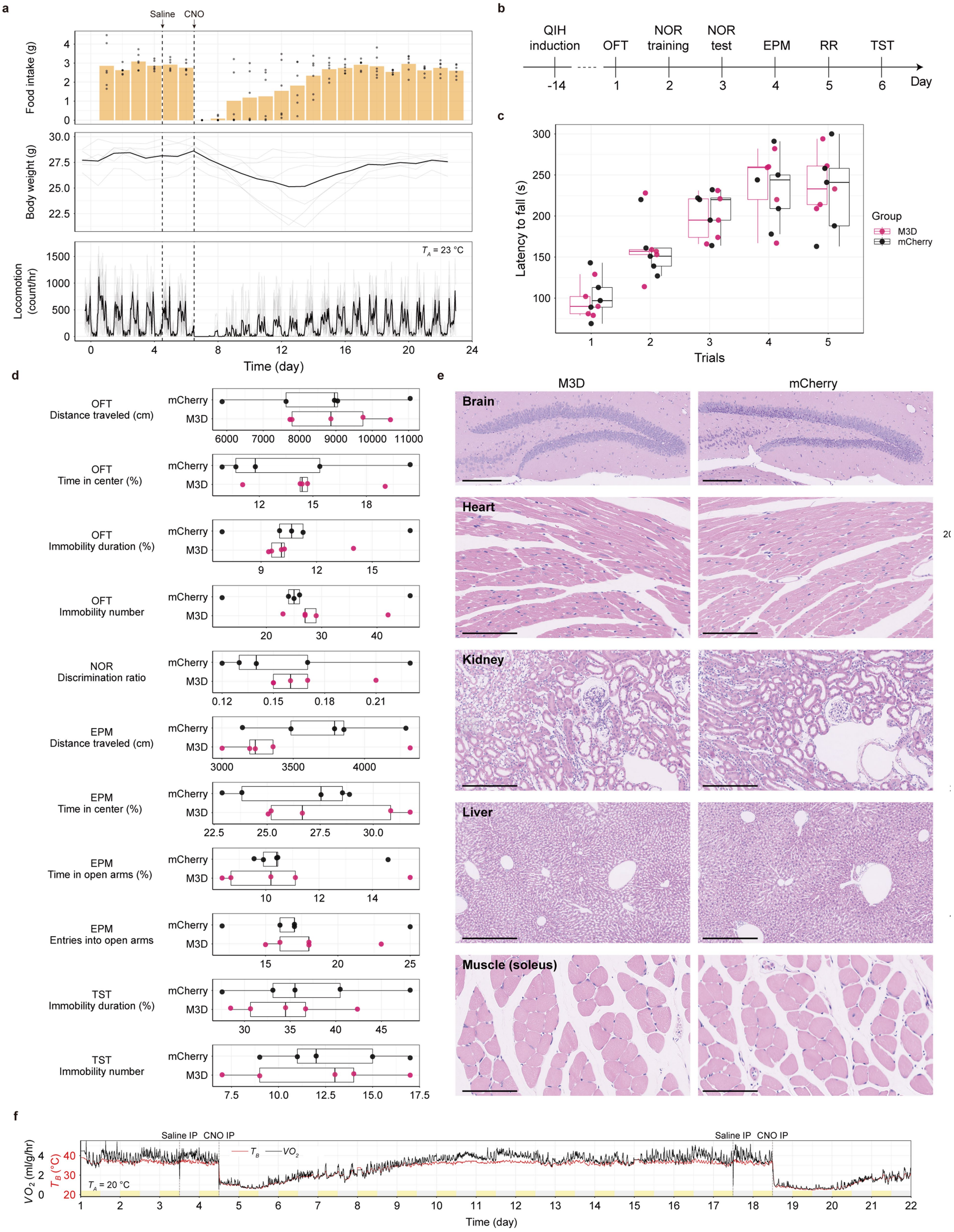
Extended Data Fig. 3 | DREADD-mediated excitation of Q neurons. a, *Qrfp* mRNA is expressed in mCherry-positive neurons in Q-hM3D mice. Dual-colour in situ hybridization for *Qrfp* and *mCherry* mRNA in brain slices prepared from Q-hM3D mice. We confirmed that CNO administration induced QIH, and subjected the mice to histological analysis. All mCherry-positive neurons were positive for *Qrfp* expression. Scale bars, 100 µm (left); 10 µm (middle, right). **b**, Representative trace of current-clamp recording from mCherry-positive Q neurons in a slice prepared from Q-hM3D mice. We performed the experiments nine times and obtained the same results. **c**, Comparison of spike frequency at baseline and after treatment with CNO ($n = 9$). **d**, Estimated distribution of spike frequency in baseline and CNO-treated slices. **e**, The estimated difference in spike frequency between CNO-treated and baseline slices was [1.44, 2.80] Hz. Because the 89% HPDI of the estimated difference is positive, the spike frequency in CNO-treated slices may be larger than baseline by more than 89%.



Extended Data Fig. 4 | T_{BAT} decreases concomitantly with body temperature during QIH. Representative traces of T_{BAT} examined by thermographic camera (orange) and body temperature measured by telemetry sensor (red) before and after induction of QIH in a Q-hM3D mouse, simultaneously. Grey bars indicate locomotor activity. Note that T_{BAT} and body temperature show almost the same values both before and after induction of QIH. A.U. arbitrary units.

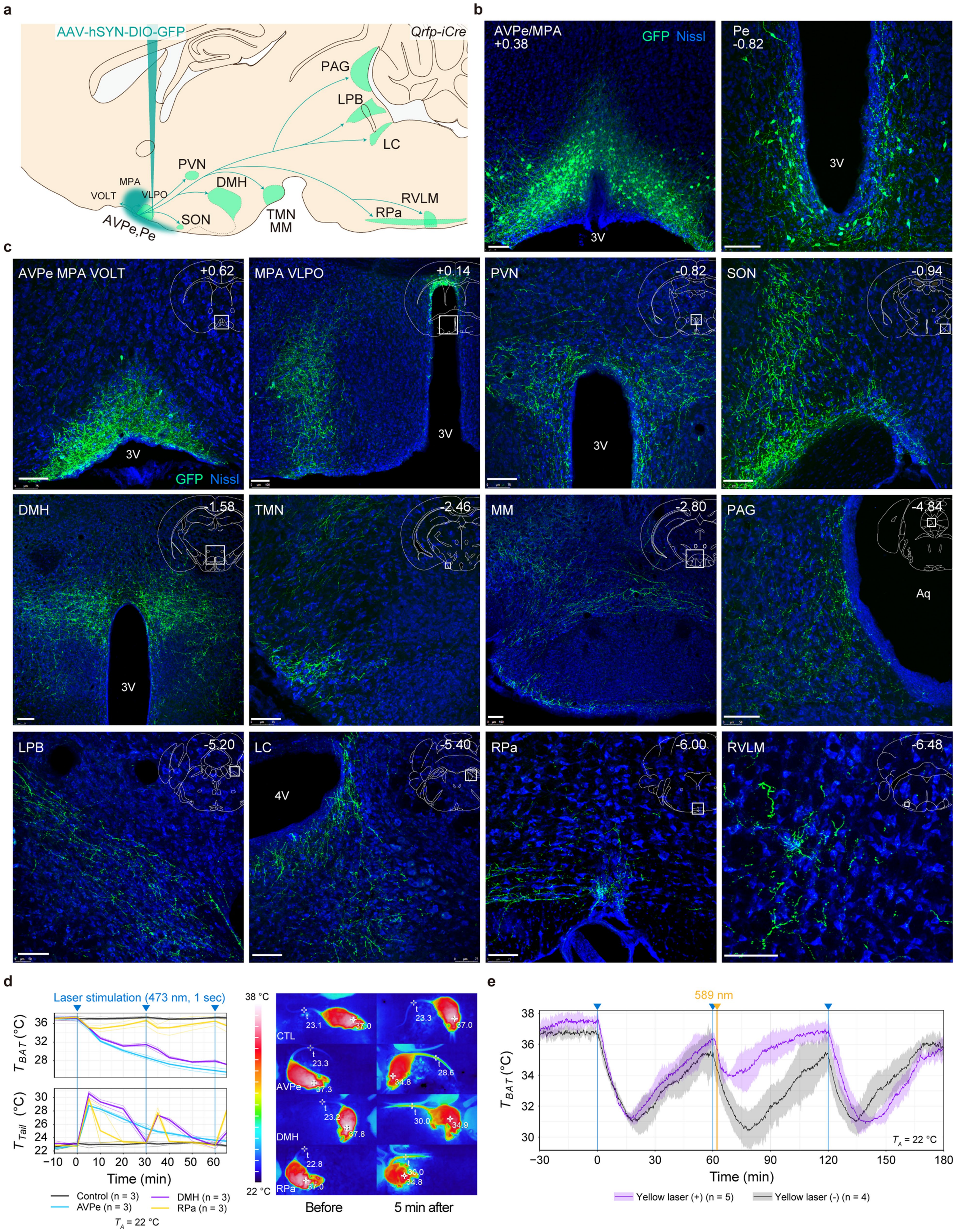


Extended Data Fig. 5 | QIH is accompanied by low heart rate, low EEG amplitude and weak respiration. ECG, EEG, VO_2 and respiratory flow were recorded during normal, FIT and QIH states ($n = 5$) in Q-hM3D mice. **a**, The one-hour median of the heart rate (HR) at minimum VO_2 during FIT was compared to that at minimum VO_2 on the day before fasting. Both VO_2 and heart rate showed marked decreases. Comparing two hours before and two hours after intraperitoneal injection of CNO, both VO_2 and heart rate were lower during QIH. The respiratory rate (RR) was undetectable in both FIT and QIH states owing to low respiratory flow. During QIH, heart rate was markedly decreased (572 and 202 beats per min, two hours before and two hours after injection of CNO, respectively). The respiratory rate of mice was reduced from 333 breaths per min to a level undetectable by the method used, suggesting that their breathing was shallow. LF and HF represent high-frequency and low-frequency power (ms^2) of HRV. **b**, Representative recordings of ECG, EEG and respiratory flow of recorded mice. Both FIT and QIH showed clear suppression of EEG amplitude. Even though movement of the chest wall was confirmed by visual inspection, respiratory flow became too low to measure the precise respiratory rate. **c**, C57BL/6J mice were fasted for 22 h from ZT0 to induce FIT ($n = 4$), followed by blood sampling at ZT22. The control group C1 ($n = 3$) was not fasted. The QIH group ($n = 6$; Q-hM3D mice) was given CNO at ZT12. Two other control groups, C2 ($n = 4$; *Qrfp*^{iCre} mice injected with AAV_{10}-DIO-mCherry into the AVPe/MPA) and C3 ($n = 4$; Q-hM3D), were injected with saline at ZT12, followed by blood sampling at ZT22. Blood glucose levels decreased during QIH, and the QIH group of mice showed hypoglycaemia and hyponatraemia compared to control groups. Both FIT and QIH groups showed high levels of ketone bodies than control groups, although the QIH group exhibited a milder phenotype than the FIT group. Levels of aspartate aminotransferase (AST), creatine kinase (CK) and potassium were lower in QIH than in FIT. ALT, alanine transaminase; GLU, glucose; LDH, lactic acid dehydrogenase; T-KB, total ketone bodies. In the box plots, the lower and upper limits of the box correspond to the first and third quartiles; the centre line denotes the median; the upper whisker extends to the largest value that is no further than 1.5 times the interquartile range (IQR); the lower whisker extends to the smallest value that is no further than 1.5 × IQR; and the dots denote observed values that are larger or smaller than the whiskers.



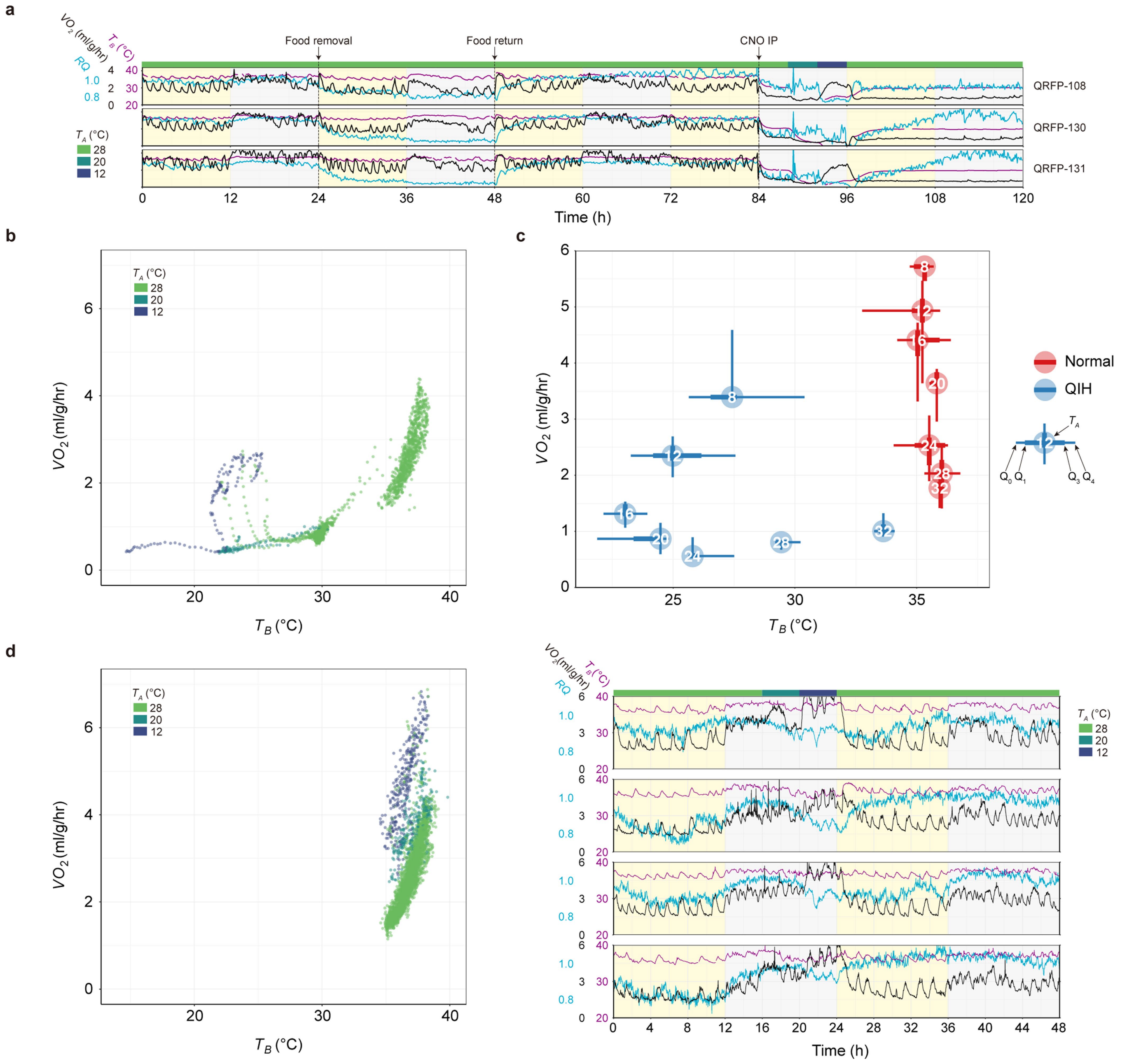
Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Mice behave normally after recovery from QIH. **a**, Food intake, body weight and activity of 6 mice were examined for 24 days before and after QIH. The first and second dashed vertical lines denote intraperitoneal injection of saline and CNO, respectively. Orange bars show the average daily food intake, and black dots represent the observed intake for each individual mouse. The bottom two panels show body weight (measured daily) and locomotor activity (measured hourly). Black lines are the average of six mice, and grey lines represent individual mice. **b**, Schematic schedule of behavioural tests. Q-hM3D mice ($n = 5$) and controls ($n = 5$; *Qrfp*^{iCre} mice with injection of AAV$_{10}$-EF1α-DIO-mCherry into the AVPe/MPA) were compared. OFT, open-field test; NOR, novel object recognition test; EPM, elevated plus maze test; RR, rotarod; TST, tail suspension test. No apparent differences were observed in any behavioural tests. **c**, Results of the rotarod test. **d**, Results of the other tests. Box plots show the distribution of each group in specific tests; all elements of the box plots are as defined in Extended Data Fig. 5. **e**, Histology of tissues before and after QIH. We histologically examined whole regions in the brain, heart, kidney, liver and soleus muscles prepared from mice that did or did not experience QIH. Tissue sections were stained with haematoxylin and eosin. No gross pathophysiological changes were apparent in any of the tissues examined. Scale bars, 200 µm (brain and kidney); 100 µm (heart and soleus muscle); 400 µm (liver). **f**, Representative traces of body temperature and VO$_2$ during QIH, which lasts for several days and can be re-induced by another injection of CNO. Line and shading denote mean and s.d. of each group.

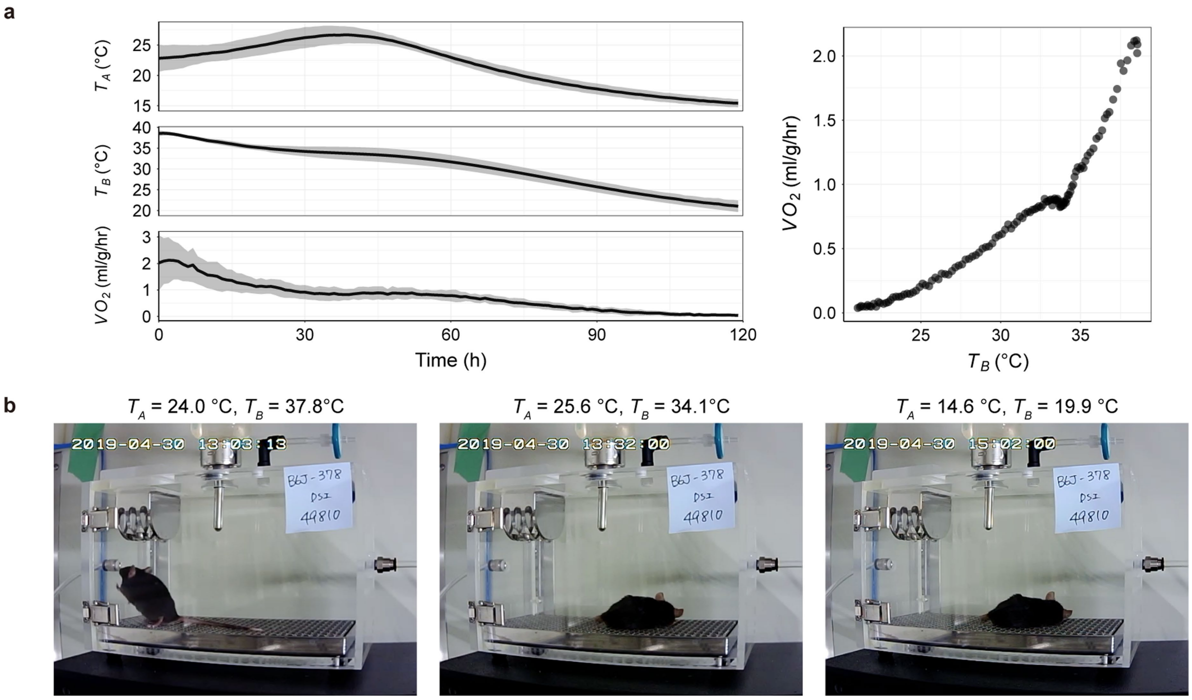


Extended Data Fig. 7 | See next page for caption.

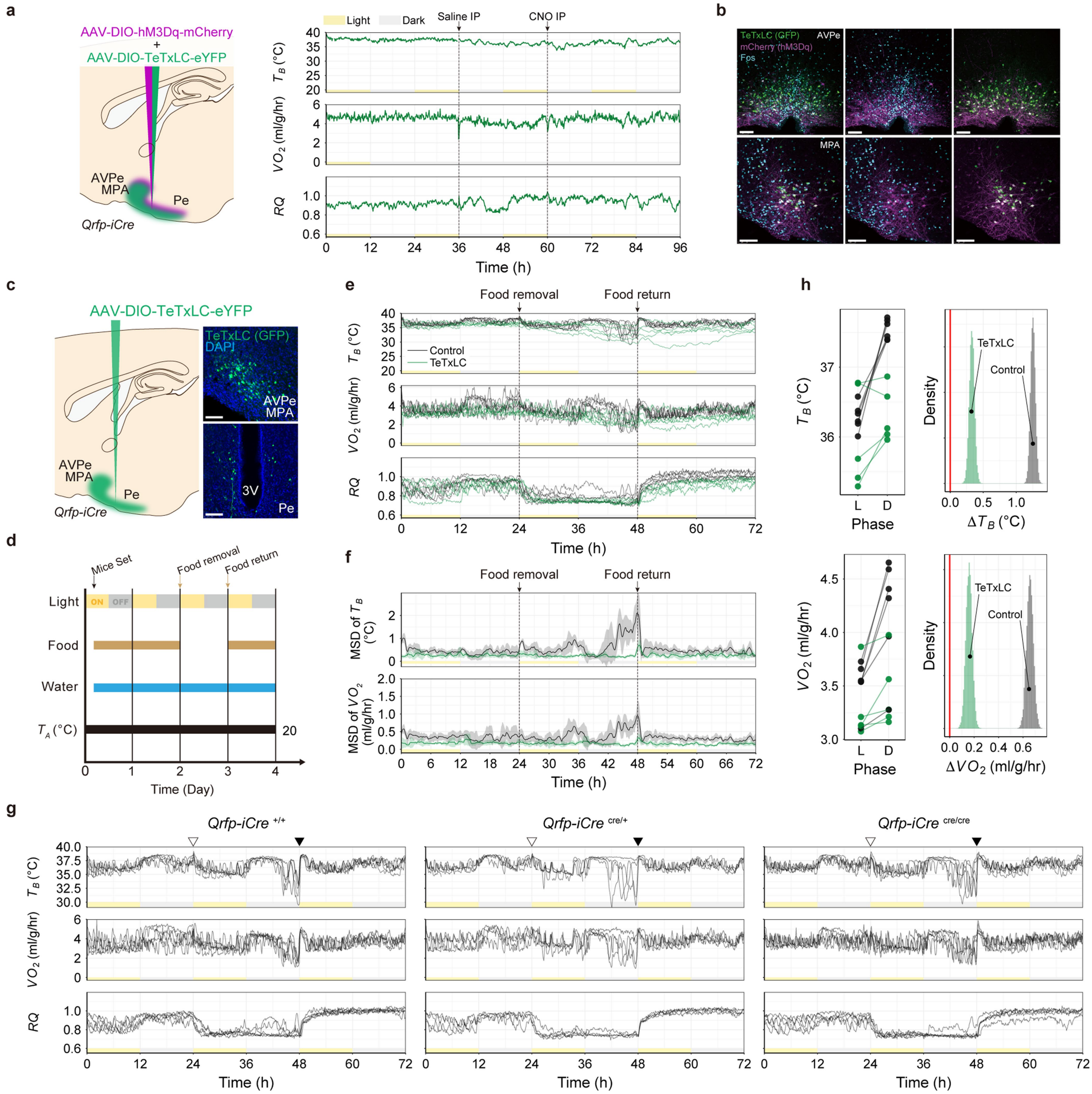
Extended Data Fig. 7 | The DMH and RPa are major target regions for the induction of QIH. a, Strategy for delineating the axonal projection patterns of Q neurons. The neurons were visualized by injecting AAV$_9$-hSYN-DIO-GFP into an anteromedial hypothalamic region of *Qrfp*^{iCre} mice to express GFP in Q neurons. **b**, Distribution of GFP-positive cell bodies of Q neurons in the AVPe/MPA and periventricular nucleus. Scale bars, 100 µm. **c**, Distribution of axons arising from Q neurons. We observed GFP-positive fibres in brain regions that are implicated in the regulation of body temperature and in sympathetic regulation. Among these regions, the DMH received especially abundant projections. Aq, aqueduct; LC, locus coeruleus; LPB, lateral parabrachial nucleus; PAG, periaqueductal grey; PVN, paraventricular hypothalamic nucleus; RVLM, rostral ventrolateral medulla; VLPO, ventrolateral preoptic area; VOLT, vascular organ of the lamina terminalis; 4V, fourth ventricle. Scale bars, 100 µm. **d**, Left, temporal changes in tail temperature of Q-SSFO mice (same mice as Fig. 2c, d) after optogenetic excitation. Right, representative images of thermographs. Optogenetic focal stimulation of Q neuron axons in the RPa also induced tail vasodilation. **e**, We implanted optic fibres in the DMH of Q-SSFO mice, applied a blue laser (1-s duration) to induce QIH and then deactivated SSFO using a 589-nm yellow laser (5-s duration) to see the effect on T_{BAT}. The first shot of blue laser in DMH fibres rapidly triggers hypothermia. A sequential shot of yellow laser 3 min after the second shot of blue laser rapidly reverses the effect of the blue laser. Because deactivation of SSFO is not propagated along axons, this further supports the importance of the DMH projections of Q neurons in the induction of QIH. Lines and shading in **d**, **e** denote mean and s.d. of each group.



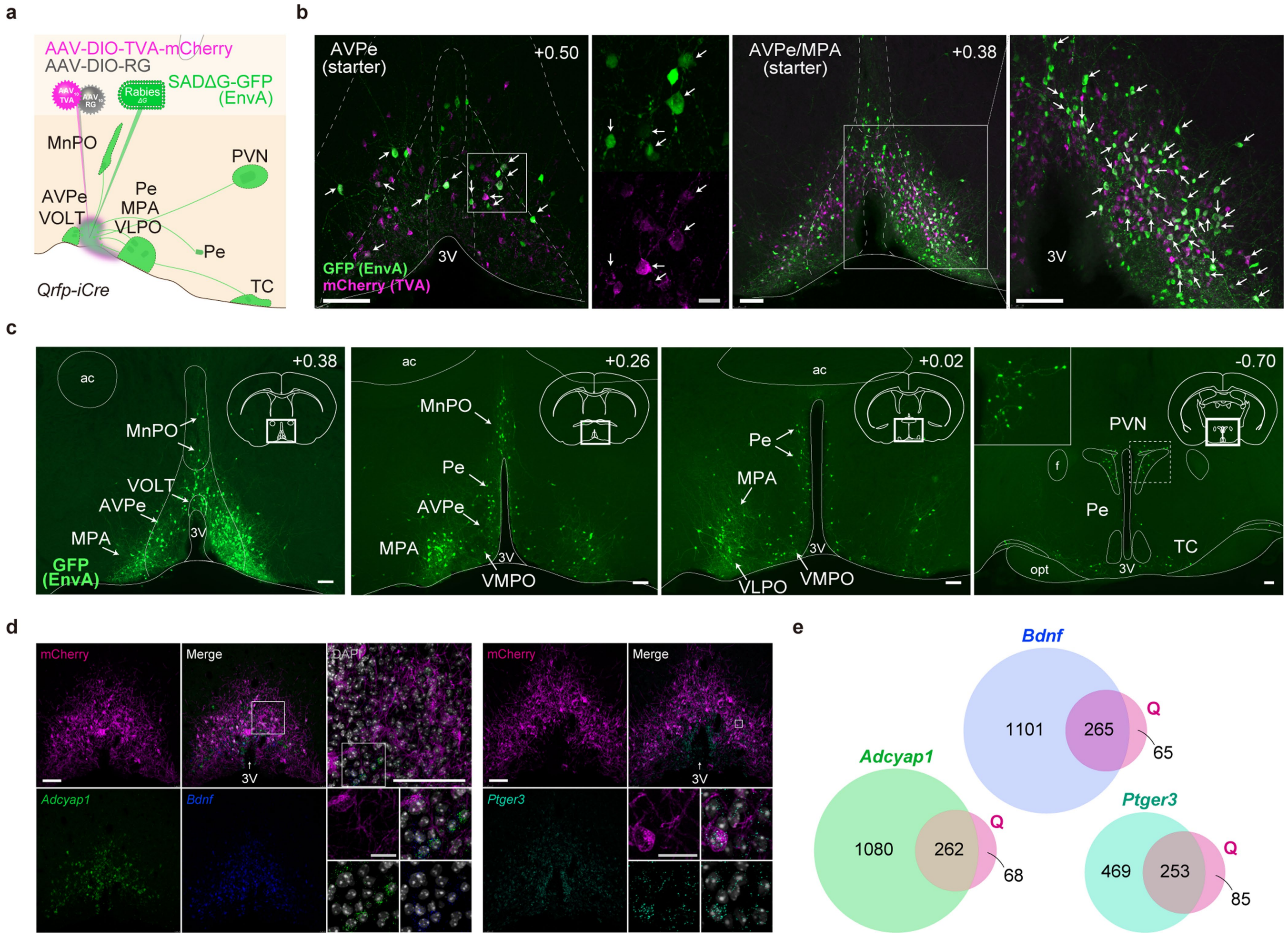
Extended Data Fig. 8 | Dynamics of set-point temperature in QIH. **a**, Transitions of metabolism when the ambient temperature was changed during QIH. See Fig. 3l for details. During QIH, when the ambient temperature was lowered from 28 °C to 20 °C, all mice showed decreased VO_2 and body temperature. By contrast, when the ambient temperature was lowered from 20 °C to 12 °C during QIH, three out of four mice showed increased VO_2 with a relatively stable body temperature. One mouse did not show an increase in VO_2, which indicates individual variance in the reduction of T_R. We confirmed that all mice spontaneously recovered from QIH. **b**, The relationship between body temperature and VO_2 during QIH with changing ambient temperature. The last 48 hours of data from Fig. 3l and **a** were merged. The colours of the dots correspond to different ambient temperatures. **c**, The relationship between the minimum body temperature and VO_2 during normal and QIH states. Data from Fig. 3b are summarized. Numbers in the dots denote the ambient temperature (°C) and the bars denote the distribution. **d**, To evaluate metabolic regulation in a normal state, wild-type C57BL/6J mice were subjected to changes in the ambient temperature. Left, the relationship between body temperature and VO_2 in all mice. Of note, body temperature is tightly controlled within a narrow range—in contrast to during QIH (**b**). Right, change in body temperature (purple), VO_2 (black) and respiratory quotient (blue) for each mouse throughout the experiment. Starting from 28 °C, the ambient temperature was lowered to 12 °C and returned to 28 °C, as shown at the top.



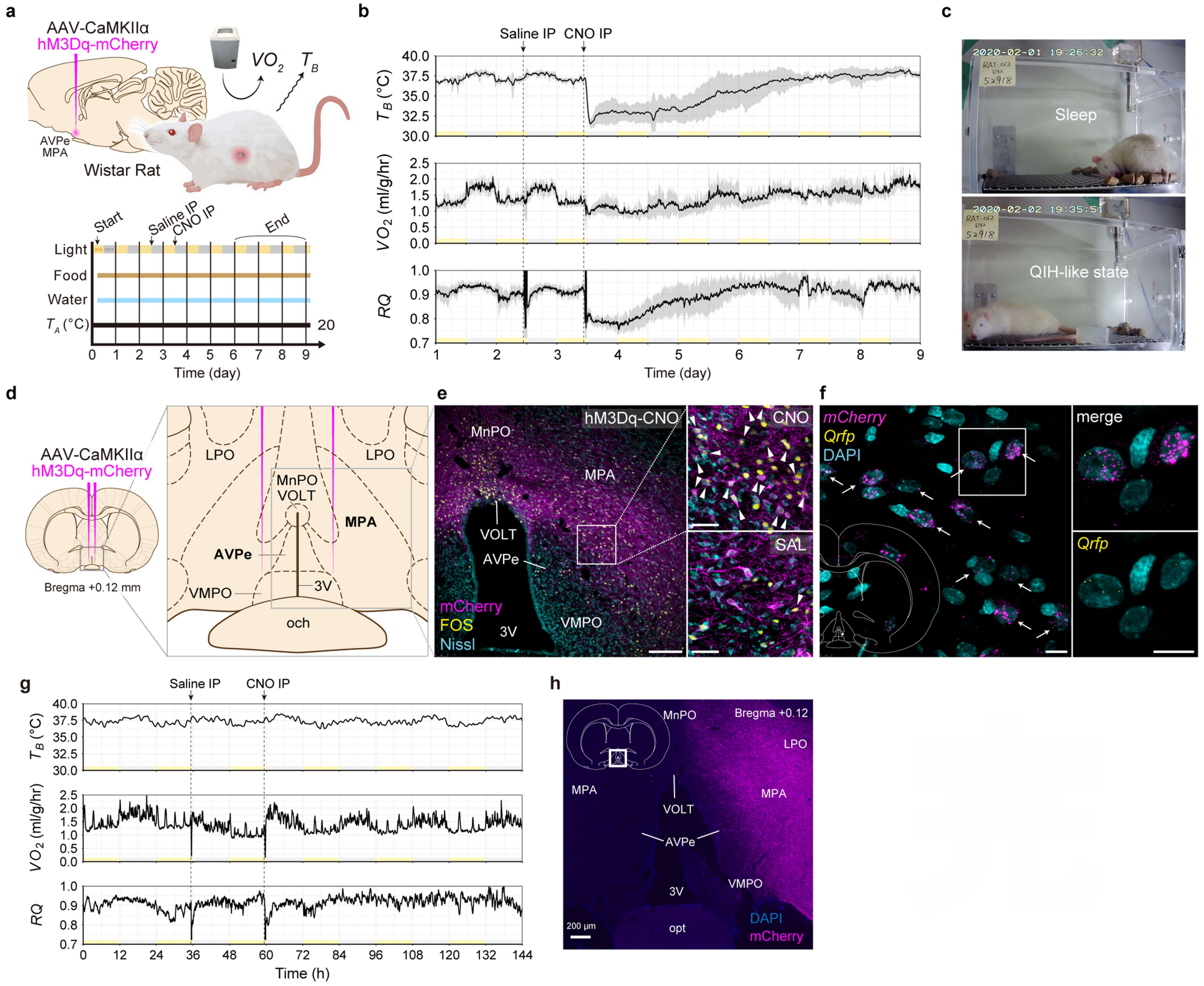
Extended Data Fig. 9 | Hypometabolism that is induced by general anaesthesia is not regulated. **a**, To evaluate how metabolic regulation during general anaesthesia was affected by ambient temperature, C57BL/6J mice ($n = 4$) were anaesthetized with 1% isoflurane at different ambient temperatures. Left (top row), the transition in ambient temperature. Starting from 28 °C, the set-point temperature of the chamber was lowered to 12 °C after 30 min. Because the anaesthetic machine was outside the experimental chamber and therefore the temperature of the anaesthetic gas was independent of that of the chamber, there was a delay in reaching the chamber set-point temperature. Left (middle and bottom rows), the transition in body temperature and VO_2. Both decrease along with the decrease in ambient temperature. Line and shading denote mean and s.d. Right, the relationship between body temperature and VO_2 in all mice. VO_2 did not increase in anaesthetized mice even at low body temperature, in contrast to in QIH (compare to Extended Data Fig. 8b). **b**, Representative postures of mice during anaesthesia. Left, the start of isoflurane inhalation. Middle, the start of the lowering of ambient temperature. Right, 90 minutes after the set-point temperature was lowered from 28 °C to 12 °C. No change in posture was seen even at extremely low body temperature.



Extended Data Fig. 10 | Blocking SNARE-complex-mediated neurotransmission in Q neurons impairs daily torpor and QIH. a, CNO had almost no effect on body temperature and VO_2 in a *Qrfp*^{iCre} mouse that was co-injected with AAV_9-DIO-hSYN-TeTxLC-eYFP and AAV_{10}-DIO-hM3Dq-mCherry into the AVPe/MPA ($n = 1$). This suggests that SNARE-mediated neurotransmission in Q neurons is indispensable for inducing QIH. **b**, Expression of TeTxLC–GFP in mCherry-positive neurons (Q neurons), shown by immunostaining 90 min after administration of CNO. Scale bar, 100 μm. **c**, Strategy for suppressing the function of Q neurons. Images show expression of TeTxLC–eYFP in the AVPe/MPA and periventricular nucleus. Scale bar, 100 μm. **d**, Schematic of FIT experiment schedule. **e**, FIT was disrupted by expressing TeTxLC in Q neurons ($n = 6$ mice for control and $n = 5$ mice for TeTxLC). The normal architecture of FIT was disrupted when neurotransmission of Q neurons was blocked in Q-TeTxLC mice. Rapid oscillatory fluctuations in metabolism were never seen in these mice. Notably, the gradual decrease in body temperature observed in these mice implies the existence of a Q-neuron-independent mechanism of metabolism reduction during FIT. **f**, The moving standard deviation (MSD; mean ± s.d.) was visualized for body temperature and VO_2 (from **e**). The low MSDs that are seen in the TeTxLC group during the fasting periods demonstrate the smaller fluctuation in this group. **g**, FIT was induced in control, *Qrfp*^{iCre} heterozygous and *Qrfp*^{iCre} homozygous mice, showing that the lack of QRFP peptide did not affect FIT. These observations suggest that Q neurons—but not QRFP—are an indispensable component in the induction of daily torpor, and have an important role in rapidly shifting body temperature during daily torpor. The open and closed triangles denote food removal and return, respectively. **h**, Silencing of Q neuron neurotransmission resulted in decreased circadian fluctuations in both body temperature and VO_2. The data from the first 24 h in panel **e** were divided into light (L) phase and dark (D) phase. Estimated differences in light phase and dark phase for both body temperature and VO_2 are shown as histograms of posterior distributions. Both body temperature and VO_2 showed higher values in the dark phase than in the light phase because posterior distributions are mostly positive. Although the TeTxLC group showed positive posterior distributions as well, the differences between dark phase and light phase were smaller than those in the control group. This suggests that the TeTxLC group had smaller circadian fluctuations in metabolism.



Extended Data Fig. 11 | Characteristics of Q neurons. To elucidate the possible neuronal mechanism that regulates the activity of Q neurons, we identified upstream neuronal populations that make direct synaptic contact with Q neurons by recombinant pseudotyped rabies virus vector (SADΔG(EnvA))-mediated labelling[45]. **a**, Procedure for visualizing input neurons that make mono-synaptic contact with Q neurons, using a rabies virus vector. After expressing TVA–mCherry and rabies glycoprotein (RG) in Q neurons using Cre-activatable AAV vectors[35] in $Qrfp^{iCre}$ mice, we injected SADΔG-GFP(EnvA) into the AVPe/MPA. **b**, Distribution of input neurons of Q neurons. Arrows show starter cells. **c**, Brain regions that contain input neurons. Input neurons were also observed in regions in and around the AVPe and periventricular nucleus, suggesting that local interneurons exist that regulate the function of Q neurons, and also indicating that Q neurons might form microcircuitry with interneurons within the AVPe/MPA and periventricular nucleus. Our results suggest that Q neurons receive relatively sparse direct inputs from intra-hypothalamic regions. As the MPA is implicated in the regulation of body temperature[46,47], reciprocal interaction between Q neurons and the MPA might have a key role in thermoregulation. **d**, In situ hybridization in neurons immunostained for mCherry in the AVPe/MPA of Q-hM3D mice. Left, expression of *Adcyap* and *Bdnf* in Q neurons; right, expression of *Ptger3* in Q neurons. **e**, Proportions of *Adcyap*-, *Bdnf*- and *Ptger3*-positive cells in Q neurons, indicating the extent to which the Q neurons overlap with genetic markers associated with thermoregulation. Numbers show the cell counts with positive signals (two mice; three slices per mouse). In the AVPe/MPA, mCherry-negative (non-Q) *Adcyap1*- and *Bdnf*-positive neurons were intermingled with Q neurons. Almost a quarter of *Adcyap1*- and *Bdnf*-positive neurons were Q neurons. We also found a small number of Q neurons that were negative for *Adcyap1* and *Bdnf*. These observations suggest that many Q neurons constitute a subpopulation of BDNF/PACAP neurons. Although a lot of Q neurons express *Adcyap1* and *Bdnf*, a previous report suggested that the warmth-sensing BDNF/PACAP neurons in the ventromedial preoptic area that project to the DMH are GABAergic[32]. As we found that excitatory Q neurons have a major role in inducing QIH (Fig. 4), Q neurons apparently constitute a unique, previously unidentified population among the group of preoptic-area neurons that are involved in thermoregulation. Notably, we found that many Q neurons express both *Vgat* and *Vglut2* (Q_H neurons) (Fig. 4a). This is consistent with a previous study reporting that many BDNF/PACAP neurons in the preoptic area express both *Vgat* and *Vglut2*[24], because Q neurons are a subset of BDNF/PACAP neurons. Prostaglandin EP3 receptor (*Ptegr3*), which is implicated in causing fever[16,48], is expressed in Q neurons. Again, the number of *Ptegr3*-positive neurons was larger than that of Q neurons, but three quarters of Q neurons expressed *Ptegr3*. This suggests that PGE2 inhibits Q neurons through acting on EP3 in Q neurons, although our inhibitory DREADD experiments did not show any effects on T_{BAT} (Extended Data Fig. 2c). ac, anterior commissure; f, fornix; MnPO, median preoptic area; opt, optic tract; VLPO, ventrolateral preoptic area; VMPO, ventromedial preoptic area.



Extended Data Fig. 12 | Induction of a QIH-like state in rats. a, Procedure for the metabolic analysis with chemogenetic activation of AVPe/MPA neurons in rats. Saline and CNO were administered just before the beginning of the dark phase. Recordings were taken until the metabolism recovered to baseline levels. **b**, Activating AVPe/MPA neurons, including Q neurons in rats, induced a QIH-like state of hypothermia and hypometabolism ($n = 7$). The lines and shadings denote mean and s.d. Body temperature, VO_2 and the respiratory quotient remained low for more than 24 h after intraperitoneal injection of CNO, as in mice during QIH, and then spontaneously returned to normal states. **c**, Representative images showing the typical posture of rats during a QIH-like state compared with during sleep. **d**, Schematic drawing of virus (AAV_{10}-CaMKIIα-hM3Dq-mCherry) injections into the AVPe/MPA of the rat brain. Stereotaxic brain maps are based on Paxinos and Watson's atlas[39]. The grey rectangular region in the right panel shows the area in which the following histological evaluations are focused. **e**, Distribution of hM3Dq-mCherry-expressing neurons in the AVPe/MPA. Arrowheads indicate hM3Dq-expressing neurons that are positive for FOS immunofluorescence 90 min after intraperitoneal injection with CNO or saline. Scale bars, 200 μm (left), 50 μm (right). **f**, *Qrfp* and *mCherry* transcripts detected in the AVPe/MPA of rats. Arrows denote co-expression of *Qrfp* and *mCherry* mRNAs. Scale bars, 10 μm. **g**, Body temperature, VO_2 and respiratory quotient before and after CNO injection in rat no. 014, which did not show a QIH-like state. **h**, Expression of hM3Dq-mCherry in the AVPe/MPA region of rat no. 014. We observed unilateral expression of hM3Dq-mCherry in the MPA region. This suggests that bilateral proper expression of hM3Dq in the AVPe/MPA is necessary to evoke the QIH-like state. Collectively, seven out of eight rats showed a QIH-like state, characterized by a prominent decrease in body temperature. In these rats, the reduction in body temperature was accompanied by a decrease in VO_2, a lowered respiratory quotient and an extended posture, showing further similarity with the QIH state in mice. The efficiency of induction of a QIH-like state in these rats is likely to be lower than that in Q-hM3D mice, owing to ectopic expression of hM3Dq in non-Q neurons within and around the AVPe/MPA.

nature research

Corresponding author(s): Takeshi Sakurai, Genshiro A. Sunagawa

Last updated by author(s): Dec 10, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a | Confirmed

☐ ☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement

☐ ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly

☒ ☐ The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.

☒ ☐ A description of all covariates tested

☒ ☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons

☐ ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)

☒ ☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted
Give P values as exact values whenever suitable.

☐ ☒ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings

☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes

☒ ☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	InfReC Analyzer NS9500 Professional software (NIPPON AVIONICS); Ponemah Physiology Platform (version 6.30, DSI); AD converter (Biotex); mass spectrometer(ARCO-2000 , ARCO System); LAS X 3.1.1.157512 for Leica TCS SP8 confocal microscope; ZEN 2.3 for Zeiss Axio Zoom.V16 microscope.
Data analysis	Every analysis and chart output was produced by R (version 3.52 or above). Every script for Bayesian inference is available on web.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Included in main text.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical method was used to determine the sample size for the animal experiments. The number of animals recorded in one group (i.e. QIH induction at 24 °C of ambient temperature) was chose based on previous experience and standards in this field. Sample sizes were described in the manuscript.
Data exclusions	We excluded mice heavier than 34 g in daily torpor experiments (Extended Data Fig.8e-h) because we found that mice weighing over 34 g did not reproducibly exhibit FIT. This criteria was not pre-specified. In the QIH-like state induction in rats (Extended Data Fig.12b), we excluded a rat which did not show expression of mCherry in the AVPe/MPA bilaterally (the data is shown in Extended Data Fig.12g, h).
Replication	See "Statistics and reproducibility"
Randomization	Because the animal we have used in this study were all inbred strains, we did not randomize the animals within a strain.
Blinding	Blinding was not performed in this study in animal experiments because nearly all data acquisition was automated and therefore have high objectivity. The histology study was not blinded due to simple human resource issue.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study	
☐	☒	Antibodies
☒	☐	Eukaryotic cell lines
☒	☐	Palaeontology
☐	☒	Animals and other organisms
☒	☐	Human research participants
☒	☐	Clinical data

Methods

n/a	Involved in the study	
☒	☐	ChIP-seq
☒	☐	Flow cytometry
☒	☐	MRI-based neuroimaging

Antibodies

Antibodies used	Commercially available antibodies were used. Primary antibodies used in the study were: rabbit anti-cFos (1:4000, ABE457, Millipore); goat anti-mCherry (1:15000, AB0040-200, Sicgen); rat anti-GFP (1:5000, 04404-84, Nacalai Tesque); mouse anti-TH (1:1000, sc-25269, Santa Cruz Biotechnology); mouse anti-orexin-A (1:200, sc-80263, Santa Cruz Biotechnology); and rabbit anti-MCH (1:2000, M8440, Sigma). Secondary antibodies used were: donkey anti-rabbit, goat, rat, or mouse, conjugated with Alexa 488, 594 or 647 (A21206, A21208, A11037, A11058, A31573, A31571, all 1:1000, all purchased from Invitrogen).
Validation	All antibodies were commercial in origin. Validation statements can be found on the manufacturer's website as following. Rabbit anti-cFos (ABE457, Millipore): https://www.merckmillipore.com/JP/ja/product/Anti-c-Fos-Antibody,MM_NF-ABE457 Goat anti-mCherry (AB0040-200, Sicgen): https://www.labome.com/product/SICGEN/AB0040-200.html Rat anti-GFP (04404-84, Nacalai Tesque): https://www.labome.com/product/Nacalai-Tesque/04404-84.html Mouse anti-TH (sc-25269, Santa Cruz Biotechnology): https://www.scbt.com/p/th-antibody-f-11 Mouse anti-orexin-A (sc-80263, Santa Cruz Biotechnology): https://www.scbt.com/ja/p/orexin-a-antibody-kk09 Rabbit anti-MCH (M8440, Sigma): https://www.sigmaaldrich.com/catalog/product/sigma/m8440?lang=ja&region=JP In the present study, validation was performed in our laboratory, testing on mouse and/or rat brain tissues.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	We used C57BL/6J mice and Wister rats. Except during torpor-inducing experiments, animals were given food and water ad libitum and maintained at TA of 23 °C at IIIS and 22 °C at BDR, relative humidity of 50%, with a 12-hr light/12-hr dark cycle. Qrfp-iCre mice were generated by homologous recombination in C57BL/6N embryonic stem cells and implantation in 8-cell-stage embryos (ICR). A targeting vector was designed to replace the entire coding region of the prepro-Qrfp sequence in exon 2 of the Qrfp gene with iCre and pgk-Neo cassette so that the endogenous Qrfp promoter drives expression of iCre (Extended Data Fig. 1). Chimeric mice were crossed with C57BL/6J females (Jackson Labs). The Pgk-Neo cassette was deleted by crossing them with FLP66 mice, which had been backcrossed to C57BL/6J mice at least 10 times. Initially, F1 hybrids from mating heterozygotes with heterozygotes were generated. We backcrossed them to C57BL/6J mice at least 8 times. All experiments were performed on iCre heterozygotes, unless indicated otherwise. Rosa26dreaddm3 and Rosa26dreaddm4 mice were generated by homologous recombination in C57BL/6N embryonic stem cells, followed by the same procedure as in Qrfp-iCre mice described above. Targeting vectors are shown in Extended Data Fig. 2a. Slc32a1tm1Lowl (referred to as Vgatflox/flox) mice and Slc17a6tm1Lowl (Vglut2flox/flox) mice were obtained from the Jackson Laboratory (Stock No: 012897 and 012898, respectively). Wistar rats were purchased from Oriental Yeast Co., Ltd. Male mice with ages ranged from 8 to 20 weeks old. Male rats were used in 8 - 11 week old.
Wild animals	No wild animal used in this study.
Field-collected samples	No samples collected at the field in this study.
Ethics oversight	All animal experiments were performed at the International Institute of Integrative Sleep Medicine (IIIS), Tsukuba University and RIKEN Center for Biosystems Dynamics Research (BDR), according to their guidelines for animal experiments. They were approved by the animal experiment committees of each institute, and thus were in accordance with NIH guidelines.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Article

Neurons that regulate mouse torpor

https://doi.org/10.1038/s41586-020-2387-5

Received: 22 January 2020

Accepted: 7 May 2020

Sinisa Hrvatin[1,6]✉, **Senmiao Sun**[1,2,6], **Oren F. Wilcox**[1], **Hanqi Yao**[1], **Aurora J. Lavin-Peter**[1], **Marcelo Cicconet**[3], **Elena G. Assad**[1], **Michaela E. Palmer**[1], **Sage Aronson**[4], **Alexander S. Banks**[5], **Eric C. Griffith**[1] & **Michael E. Greenberg**[1]✉

The advent of endothermy, which is achieved through the continuous homeostatic regulation of body temperature and metabolism[1,2], is a defining feature of mammalian and avian evolution. However, when challenged by food deprivation or harsh environmental conditions, many mammalian species initiate adaptive energy-conserving survival strategies—including torpor and hibernation—during which their body temperature decreases far below its homeostatic set-point[3–5]. How homeothermic mammals initiate and regulate these hypothermic states remains largely unknown. Here we show that entry into mouse torpor, a fasting-induced state with a greatly decreased metabolic rate and a body temperature as low as 20 °C[6], is regulated by neurons in the medial and lateral preoptic area of the hypothalamus. We show that restimulation of neurons that were activated during a previous bout of torpor is sufficient to initiate the key features of torpor, even in mice that are not calorically restricted. Among these neurons we identify a population of glutamatergic *Adcyap1*-positive cells, the activity of which accurately determines when mice naturally initiate and exit torpor, and the inhibition of which disrupts the natural process of torpor entry, maintenance and arousal. Taken together, our results reveal a specific neuronal population in the mouse hypothalamus that serves as a core regulator of torpor. This work forms a basis for the future exploration of mechanisms and circuitry that regulate extreme hypothermic and hypometabolic states, and enables genetic access to monitor, initiate, manipulate and study these ancient adaptations of homeotherm biology.

Torpor and hibernation enable warm-blooded animals to survive harsh environments that are otherwise incompatible with life[7,8]. Although constituting complex multifaceted behaviours[9–11], perhaps the most notable feature of these states is the profound decrease in core body temperature to far below its tightly controlled homeostatic set-point[3,4]. Several regions in the mammalian brain—including the preoptic area (POA)[12], the dorsomedial hypothalamus[13] and the raphe nuclei[14]—have been implicated in the coordination of temperature regulation[2]. Specific electrophysiologically and/or molecularly defined cellular components of homeostatic thermoregulation have been identified, including neurons that are sensitive to changes in ambient temperatures[15–18] and/or local brain temperature[19–21]. However, although a picture of the circuitry that underpins normal thermoregulation is beginning to emerge[1,2], how animals disengage or circumvent these conserved homeostatic mechanisms in response to environmental challenges to enter profoundly hypothermic states such as torpor and hibernation remains a central question in homeotherm biology.

To study the mechanisms that underlie the initiation of these adaptive hypothermic states, we used a model of fasting-induced torpor in laboratory mice (*Mus musculus*). Mice placed in environments that are devoid of food and are colder than their thermoneutral point (around 30 °C)[22,23] alternate between two survival strategies: high-risk food-seeking behaviour[24] and energy-conserving torpor[5,6,25]. Mouse torpor is a complex natural behaviour that is characterized by repeated bouts of profoundly reduced core body temperature (as low as 20 °C), along with decreases in movement, sensory perception, breathing, heart rate and metabolic rate[4–6,10,26–28]. To study fasting-induced torpor, mice were housed at 22 °C and implanted with telemetric temperature probes. Whereas fed mice maintained a core body temperature (T_b) higher than 35.1 ± 0.2 °C, all mice that were food-restricted for 24 hours experienced one or more bouts of torpor, which we characterized as a precipitous drop in core body temperature (greater than 1 °C per 20 min), a period of deep hypothermia (T_b of 24–35 °C) lasting up to several hours, and finally, arousal from torpor (Fig. 1a). Each bout of torpor was preceded by a 45.1 ± 4.6% decrease in metabolic rate and was accompanied by reduced movement (73.4 ± 7.0% reduction; Fig. 1a–c, Extended Data Fig. 1a). Although circadian rhythms[28], leptin signalling[27,29], sympathetic nervous system activity and adipose tissue thermogenesis[26,30] have all been shown to modulate torpor, the mechanisms by which animals trigger and regulate this natural hypothermic state remain unknown.

---

[1]Department of Neurobiology, Harvard Medical School, Boston, MA, USA. [2]Program in Neuroscience, Harvard Medical School, Boston, MA, USA. [3]Image and Data Analysis Core, Harvard Medical School, Boston, MA, USA. [4]Neurophotometrics, Ltd., San Diego, CA, USA. [5]Division of Endocrinology, Diabetes and Metabolism, Beth Israel Deaconess Medical Center, Boston, MA, USA. [6]These authors contributed equally: Sinisa Hrvatin, Senmiao Sun. ✉e-mail: sinisa_hrvatin@hms.harvard.edu; michael_greenberg@hms.harvard.edu

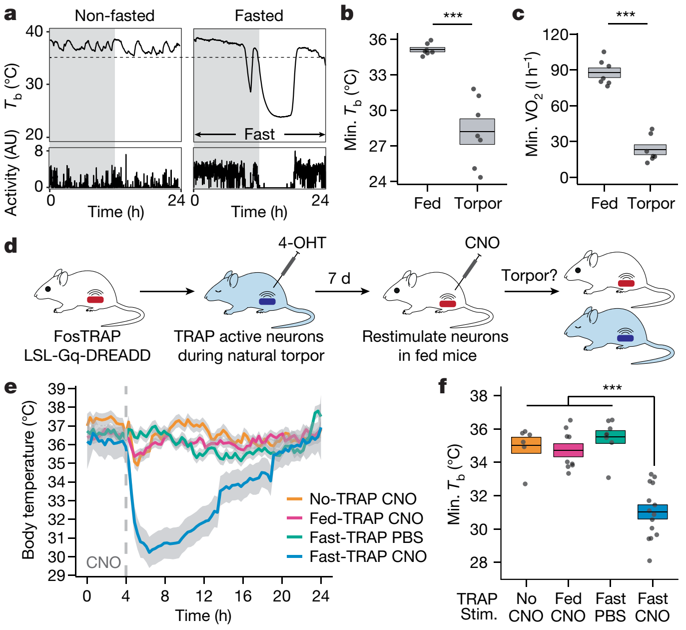


Fig. 1 | Neuronal activity induces key features of torpor. a, Core body temperature (T_b) and gross motor activity of a representative non-fasted (left) or fasted (right) mouse over 24 h. Fasted mice enter torpor, whereas non-fasted mice do not. The grey and white backgrounds indicate 12-h periods of darkness and light, respectively. The dashed line indicates the minimum T_b observed in non-fasted mice. **b**, **c**, Minimum T_b (**b**) and the minimum metabolic rate as measured by the volume of oxygen consumed (VO_2) (**c**) in non-fasted (fed) and fasted (torpor) mice ($n = 7$ mice, ***$P = 6 \times 10^{-4}$ (**b**), ***$P = 1 \times 10^{-5}$ (**c**)). **d**, Schematic showing the procedure for gaining genetic control over torpor-regulating neurons. Neurons active during torpor in FosTRAP, LSL-Gq-DREADD mice are TRAPed by 4-OHT administration and chemogenetically restimulated 7 days later in non-fasted mice by treatment with CNO. **e**, CNO-induced reactivation of 4-OHT-TRAPed neurons that are active during torpor entry triggers a decrease in T_b characteristic of mouse torpor (fast-TRAP CNO, $n = 14$ mice). The same mice injected with PBS (fast-TRAP PBS, $n = 8$ mice), or control mice in which neurons were not TRAPed (no-TRAP CNO, $n = 6$ mice) or were TRAPed during a non-torpid state (fed-TRAP CNO, $n = 9$ mice), did not show a decrease in T_b upon CNO administration. The dashed line indicates the onset of CNO or PBS administration, shading indicates 95% confidence interval. **f**, The minimum T_b after CNO administration is lower in fast-TRAP CNO ($n = 14$) mice compared to no-TRAP CNO ($n = 6$, $P = 6.2 \times 10^{-4}$), fed-TRAP CNO ($P = 2.4 \times 10^{-6}$) or fast-TRAP PBS ($P = 2.5 \times 10^{-5}$) mice. For the box plots, the centre line and box boundaries indicate mean ± s.e.m. *P* values were calculated using a two-tailed Mann–Whitney *U*-test, ***$P < 0.001$.

Torpor-associated circuit activity

In principle, entry into torpor could be triggered by circulating factors capable of reducing metabolic rate and/or by changes in thermoregulatory neural circuit activity. Consistent with the idea that altered circuit activity contributes to entry into torpor[26], staining for FOS—a marker of neuronal activity-induced transcription[31]—followed by whole-brain imaging and machine learning-enabled registration of the FOS signal to the Allen Mouse Brain Atlas[32] (Methods) revealed several brain regions that are active during fasting-induced torpor. As might be expected, we observed neuronal activity in brain regions that regulate hunger, feeding and energy balance[33,34], as well as in thermoregulatory areas[1,2,35] and in a large number of other brain regions. This finding that brain circuits are engaged as fasted mice enter torpor suggests that these circuits might potentially drive the entry process (Extended Data Fig. 1b–f).

To determine whether neural circuit activity is sufficient to induce torpor phenotypes independent of caloric restriction, we used genetic tools that enable the expression of a chemically activated receptor, Gq-DREADD (Gq-coupled Designer Receptor Exclusively Activated by Designer Drug), specifically in the neurons that are active as mice enter torpor. Reactivation of the putative torpor-regulating neurons by administration of the Gq-DREADD-activating synthetic ligand clozapine *N*-oxide (CNO) to the mice enabled us to determine whether the reactivation of these neurons alone—without caloric restriction—is sufficient to induce torpor-associated phenotypes. For this experiment, we used mice harbouring a tamoxifen-dependent form of Cre recombinase driven from the *Fos* locus ($Fos^{2A\text{-}iCreERT2}$, TRAP2[36]) together with an allele of the Gq-coupled receptor that is expressed in a Cre-dependent manner (R26-LSL-Gq-DREADD[37]). When these 'FosTRAP-Gq' mice are fasted to induce entry into torpor, the neurons that are active—and thus potentially mediate torpor entry—induce FOS and CreERT2. When these mice are exposed to 4-hydroxytamoxifen (4-OHT), the CreERT2 recombines the R26-LSL-Gq-DREADD allele, leading to the persistent expression of Gq-DREADD and enabling these specific neurons—referred to hereafter as 'TRAPed' neurons—to be activated at a later time by the administration of CNO (Fig. 1d).

FosTRAP-Gq mice ($n = 14$) were fasted and 4-OHT was administered as they entered natural torpor. After several days of recovery from fasting, these mice were administered the DREADD-activating ligand CNO to chemogenetically restimulate the neurons that were TRAPed during natural torpor (Fig. 1d, Methods). Notably, we found that the stimulation of neurons that were previously active during fasting and torpor was sufficient to induce the robust decrease in core body temperature and locomotor activity associated with natural torpor, despite the absence of caloric restriction. This effect was dependent on CNO administration and on previous 4-OHT-mediated TRAPing in the fasted state (Fig. 1e, f, Extended Data Fig. 1g, h). Although we cannot exclude a contribution either from fasting-regulated neurons that are active before or after torpor or from non-neuronal cells, this result suggests that the systemic recapitulation of torpor-associated neuronal circuit activity is sufficient to acutely induce key behavioural and physiological features of torpor.

avMLPA neurons regulate features of torpor

To identify the brain areas that were labelled using the TRAP approach, we immunostained brain sections of these FosTRAP-Gq mice for the haemagglutinin-tagged Gq-DREADD protein. Whole-brain imaging revealed widespread expression of the Gq-DREADD protein, with 190 differentially labelled regions identified between mice TRAPed in a fasted state (fast-TRAPed) compared with mice in a fed state (fed-TRAPed; Extended Data Fig. 1i–k, Supplementary Table 1). A strong correlation across brain regions was observed between the number of FOS^+ cells in torpid mice and the levels of Gq-DREADD expression in fast-TRAPed mice, suggesting that our TRAP approach labelled—as intended—neurons that are active and induce FOS during torpor (Extended Data Fig. 1l). Although in principle the simultaneous activation of multiple neural populations across several distributed brain areas might be required to orchestrate torpor, we proposed that circuit activity within a single brain region might have a major role in the regulation of torpor. To address this possibility, we designed a screen across the brain regions that were identified by Gq-DREADD staining in fast-TRAPed mice. By stereotactic injection into FosTRAP mice, which do not express an endogenous Gq-DREADD, we administered adeno-associated viruses (AAVs) expressing Cre-dependent Gq-DREADD fused to mCherry (AAV-DIO-Gq–mCherry); this enabled TRAPing restricted to the injection area, the expression of Gq-DREADD–mCherry, and the subsequent chemogenetic restimulation of the neurons active during natural torpor within just the injected region (Fig. 2a). For these studies, we focused on the torpor-associated decrease in core body temperature. We injected FosTRAP mice ($n = 54$) in different areas of the anterior hypothalamus, a region of the brain involved in thermoregulation and energy balance[1,2] that

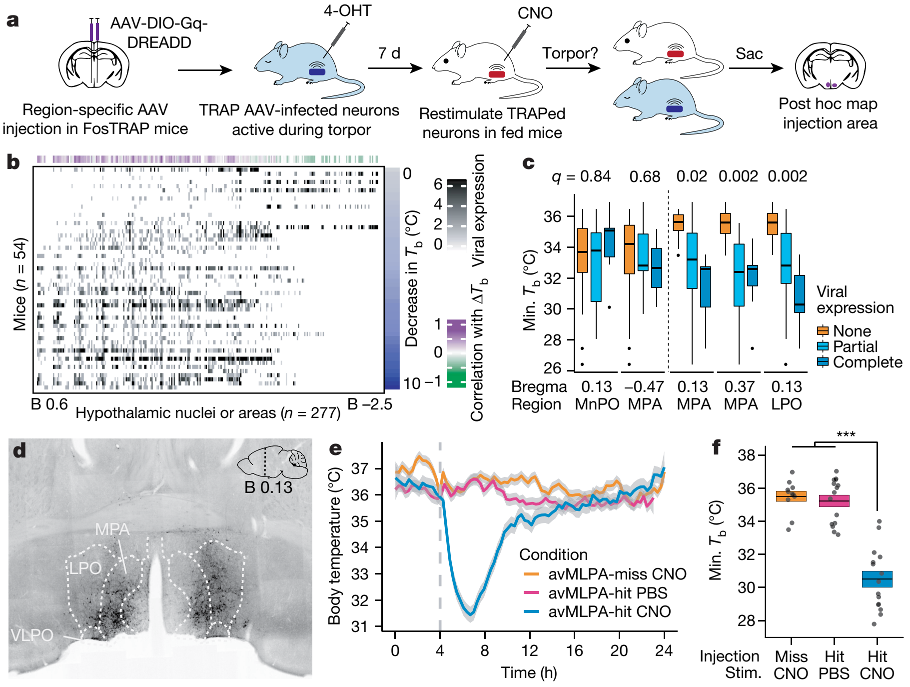


Fig. 2 | Identification of brain regions that regulate torpor. **a**, Schematic showing the procedure for identifying which hypothalamic regions contain torpor-regulating neurons. **b**, Quantification of AAV-DIO-Gq-DREADD–mCherry expression in mice TRAPed during fasting-induced torpor. Hypothalamic nuclei ($n = 277$) are plotted on the basis of their anterior–posterior (AP) coordinates relative to bregma (B). Mice ($n = 54$) are ranked on the basis of the decrease in core body temperature (ΔT_b) observed after chemogenetic stimulation of TRAPed neurons. ΔT_b is correlated with viral expression in each region (Pearson correlation). **c**, Two regions in which viral expression did not show significant correlation to ΔT_b (first two columns) and three regions in which viral expression did show significant correlation to ΔT_b (final three columns). False discovery rate (FDR)-corrected q value, Pearson correlation test, $n = 54$ mice. The minimum T_b was calculated across all mice grouped on the basis of degree of viral expression into 'none', 'partial' or 'complete'. In the box plots the centre line denotes the median, the box boundaries mark the interquartile range (IQR) and the whiskers extend to 1.5 × IQR. MnPO, median preoptic nucleus; MPA, medial preoptic area; LPO, lateral preoptic area. **d**, Representative coronal section from an avMLPA-injected mouse ($n = 15$ mice). VLPO, ventrolateral preoptic nucleus. **e**, **f**, Chemogenetic restimulation of avMLPA TRAPed neurons (avMLPA-hit CNO, $n = 15$ mice), the same mice ($n = 15$) injected with PBS (avMLPA-hit PBS), or control mice in which the avMLPA was missed (avMLPA-miss, $n = 11$ mice). **e**, A decrease in T_b characteristic of torpor is seen only in avMLPA-hit CNO mice. The dashed line indicates CNO or PBS administration, grey shading indicates 95% confidence interval of T_b. **f**, The minimum T_b after CNO administration is lower in avMLPA-hit mice compared with avMLPA-miss mice ($P = 1.8 \times 10^{-6}$) or with avMLPA-hit mice injected with PBS ($P = 5.8 \times 10^{-7}$). Two-tailed Mann–Whitney U-test, ***$P < 0.001$. For the box plots, the centre line and box boundaries indicate mean ± s.e.m.

showed substantial FOS expression during fasting-induced torpor (Extended Data Fig. 1). After recovery, these mice—hereafter denoted 'FosTRAP$^{\text{AAV-Gq}}$' mice—were fast/torpor-TRAPed, enabling persistent expression of the viral Cre-dependent Gq-DREADD–mCherry in fast- and torpor-active neurons within the virally injected region. Several days later, we administered CNO to stimulate Gq-DREADD-expressing neurons selectively in the injected region of the hypothalamus to test whether the stimulation of these neurons would result in a decrease in body temperature, as in natural torpor. The reduction of core body temperature was correlated with the anatomical expression of the virally derived Gq-DREADD–mCherry across 277 hypothalamic nuclei or areas (Fig. 2b, Methods). Leveraging the variability across injection sites in different mice ($n = 54$), this unbiased screen identified the anterior and ventral portions of the medial and lateral preoptic area (avMLPA) as key regions, with mice injected in the avMLPA exhibiting a large decrease in core body temperature in response to CNO as compared with mice in which these regions were not transduced (4.90 ± 0.68 °C compared with 0.89 ± 0.25 °C, $P = 1.8 \times 10^{-6}$; Fig. 2c–f, Extended Data Fig. 2a–c, Supplementary Tables 2, 3). FosTRAP$^{\text{AAV-Gq}}$ mice that were injected selectively in the avMLPA, TRAPed during torpor, allowed sufficient time to recover and then stimulated with CNO showed a decrease in metabolic rate of 30.4 ± 8.5% ($P < 2.4 \times 10^{-3}$) and a decrease in gross motor activity of 58.7 ± 5.2% ($P < 1.5 \times 10^{-3}$), consistent with the features observed in natural torpor (Extended Data Fig. 2d–g). Moreover, sectioning and staining brains for the virally derived Cre-dependent Gq-DREADD–mCherry fusion protein revealed projections from torpor-TRAPed avMLPA (avMLPA$^{\text{torpor}}$) neurons to the dorsomedial hypothalamus, arcuate nucleus, periaqueductal grey and raphe pallidus—regions that are known to modulate energy balance and adipose tissue thermogenesis, which are processes thought to be involved in the induction of torpor[38,39] (Extended Data Fig. 2h–j). Together, these findings identify a brain area in which the reactivation of torpor-associated neurons is sufficient to acutely induce torpor-like behavioural and physiological changes, and suggest that avMLPA$^{\text{torpor}}$ neurons may represent a critical node in the circuit that regulates natural entry into torpor.

Molecular analysis of avMLPA$^{\text{torpor}}$ neurons

The mammalian POA houses an interconnected ensemble of cell types that are involved in temperature[15–17,20], fluid[36] and cardiovascular homeostasis[40], as well as mating, parental behaviours[41] and sleep[42–44]. To catalogue the diversity of neuronal cell types present in the avMLPA, and identify which among them are active and TRAPed during torpor, we adapted a high-throughput single-nucleus RNA-sequencing (snRNA-seq)-based strategy[45] (Fig. 3a). Five FosTRAP mice were injected with AAV-DIO-Gq-DREADD–mCherry. Four of these mice were TRAPed during torpor, while one was kept as a non-TRAPed control. Their anteroventral POAs were dissected and dissociated, and 44,669 single nuclei

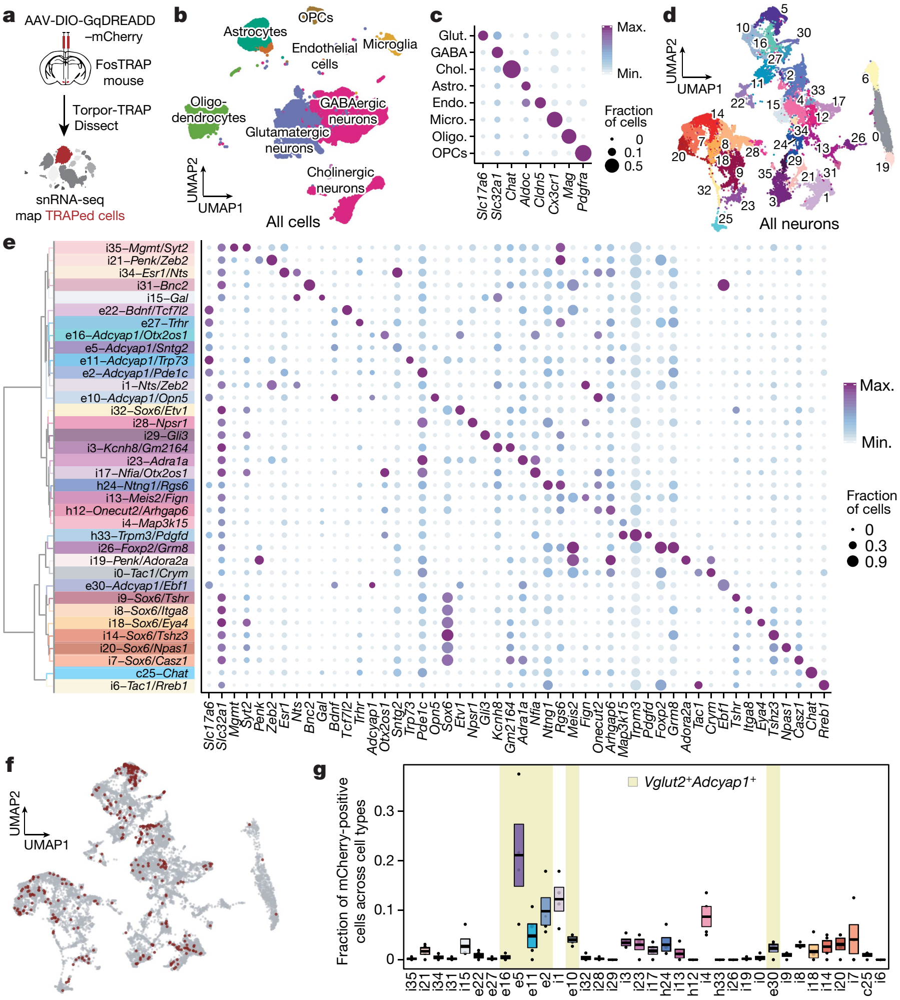


Fig. 3 | Molecular characterization of torpor-associated avMLPA neurons. **a**, Schematic showing the procedure for the molecular characterization of avMLPA$^{\text{torpor}}$ cells. AAV-DIO-Gq-DREADD is injected into the avMLPA (*n* = 5 mice). After TRAPing, the avMLPA is microdissected and analysed by snRNA-seq. **b**, Uniform manifold approximation and projection (UMAP) plot of 39,562 nuclei from the avMLPA of 5 mice. Colours group the main cell types. OPCs, oligodendrocyte precursor cells. **c**, Expression of the indicated marker genes across different cell types (named as abbreviations of the cell types in **b**). **d**, UMAP plot of 28,103 neuronal nuclei. Colours group the 36 neuronal subtypes. **e**, Expression of marker genes across different neuronal cell types. Cell types are organized on the basis of hierarchical clustering. The acronym comprises the neuronal class (e, excitatory; i, inhibitory; h, hybrid; c, cholinergic) and the cluster number, followed by select marker genes. **f**, UMAP plot of 17,424 neuronal nuclei that were transduced by the AAV (grey) and 342 neuronal nuclei that were TRAPed during torpor (red). **g**, Distribution of TRAPed neurons across all neuronal cell types. For the box plots, the centre line and box boundaries indicate mean ± s.e.m. (*n* = 4 mice). Yellow shading indicates cell types expressing *Adcyap1* and *Vglut2*. The acronym comprises the neuronal class and the cluster number. In **c**, **e**, the colour of the circle denotes the mean expression across all nuclei normalized to the highest mean across cell types, and the size of the circle represents the fraction of nuclei in which the marker gene was detected.

were analysed using snRNA-seq at an average depth of 1,286 genes and 2,083 transcripts per nucleus (Extended Data Fig. 3a–e). Unsupervised graph-based clustering delineated major neuronal and non-neuronal cell classes, and further clustering of just the neuronal subpopulation (*n* = 28,103) identified a considerable diversity of 24 GABAergic, 8 glutamatergic, 3 hybrid (GABAergic and glutamatergic) and one cholinergic neuronal cell type, consistent with the large diversity of cell types present in the POA[41] (Fig. 3b–e, Extended Data Fig. 4, Supplementary Table 4). The robustness of the obtained clusters was confirmed by subsampling analysis (Extended Data Fig. 3f). Notably, 17,424 of the 28,103 sequenced neurons—representing all 36 cell types—expressed AAV-derived transcripts, consistent with broad tropism of AAV8 in the hypothalamus.

We analysed the expression of Gq-DREADD–mCherry transcripts as a means to identify which among the transduced neurons were TRAPed during torpor. This approach detected 342 torpor-TRAPed neurons among 15,056 transduced neurons in the four TRAPed mice, and displayed a low false-positive and false-negative rate (Extended Data Fig. 5a–e, Methods). TRAPed neurons represented several avMLPA cell types, suggesting that several neuronal populations were active during fasting-induced torpor (Fig. 3f). However, the largest subset (42.6 ± 3.5%) of all torpor-TRAPed cells consisted of several populations of glutamatergic *Adcyap1*⁺ neurons (Fig. 3g, Extended Data Fig. 5f–i), a result that was subsequently confirmed using in situ hybridization methods (Extended Data Fig. 6, Extended Data Fig. 7). Differential gene expression analysis between TRAPed and non-TRAPed *Vglut2*⁺*Adcyap1*⁺ neurons identified markers of e5 neurons, consistent with preferential TRAPing of this molecularly defined subtype of *Adcyap1*⁺ neurons (Supplementary Table 5).

Stimulation of torpor-associated neurons

Together with previous work that describes distinct populations of GABAergic and glutamatergic warm-sensitive thermoregulatory

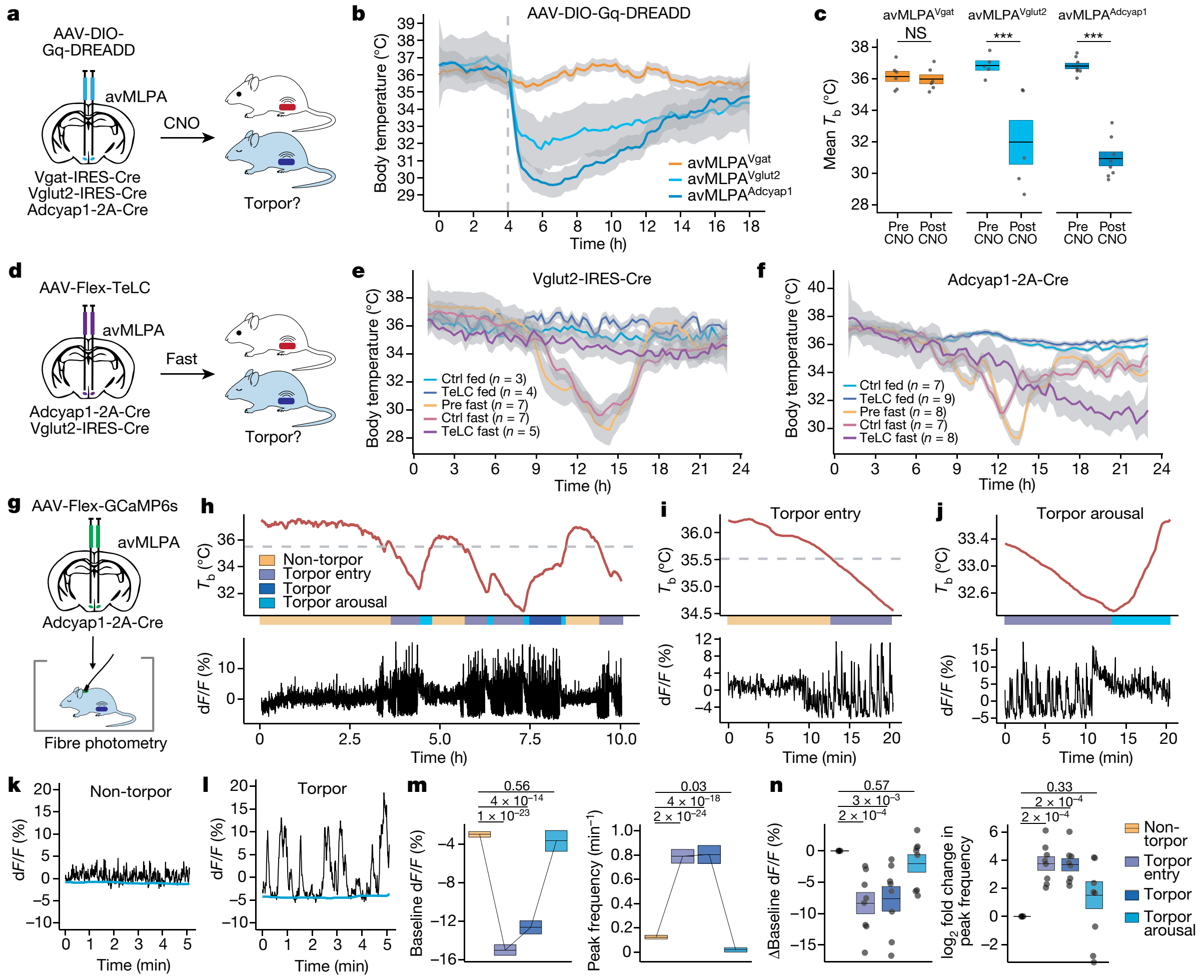


Fig. 4 | Sufficiency, necessity and natural activity of avMLPA neuronal subpopulations during torpor. a–**c**, The injection of AAV-DIO-Gq-DREADD and subsequent chemogenetic stimulation of avMLPA^{Vgat}, avMLPA^{Vglut2} and avMLPA^{Adcyap1} neurons ($n = 6$, 5 and 8 mice, respectively). **a**, Schematic showing the experimental procedure. **b**, Change in T_b after chemogenetic stimulation with CNO (indicated by the dashed line). **c**, Mean T_b before and after chemogenetic stimulation of avMLPA^{Vgat} (NS, $P = 0.48$), avMLPA^{Vglut2} (***$P = 7.9 \times 10^{-3}$), and avMLPA^{Adcyap1} neurons (***$P = 1.6 \times 10^{-4}$). **d**, Schematic showing the injection of AAV-Flex-TeLC to inhibit synaptic transmission. **e**, **f**, T_b of fed and fasted mice in which avMLPA^{Vglut2} (**e**) or avMLPA^{Adcyap1} (**f**) neurons remained un-injected (pre), were injected with a control AAV (ctrl), or were injected with AAV-Flex-TeLC (TeLC). Coloured lines indicate the mean across mice; grey shading indicates the 95% confidence interval. The number of mice for each condition is indicated in parentheses. **g**, Schematic showing the injection of AAV-Flex-GCaMP6s and fibre photometry recording from avMLPA^{Adcyap1} neurons. **h**–**j**, Recording sessions in fasted mice showing T_b and the normalized GCaMP6s signal. Coloured bars indicate the different states, classified on the basis of T_b. The dashed line indicates the transition between non-torpor and torpor states (Methods). *F* denotes the fluorescence intensity of GCaMP6s, and d*F*/*F* is calculated by dividing the smoothed calcium-dependent GCaMP6s signal with the Ca^{2+}-independent scaled fit. **h**, Example 10-h trace spanning non-torpor and torpor states. **i**, **j**, Example 20-min traces spanning torpor entry (**i**) and torpor arousal (**j**). **k**, **l**, Example photometry signals in non-torpor (**k**) and torpor (**l**) states. The baseline signal is indicated in blue. **m**, Mean baseline (left) and peak frequency (right) of the fibre photometry signal in one mouse across non-torpor and torpor states. The legend is displayed in **n**, and *P* values are indicated above the plots. **n**, Difference in average baseline (left) and $\log_2$ fold change in average peak frequency (right) of fibre photometry signal between non-torpor and torpor states ($n = 8$ mice, *P* values indicated above plots). For the box plots, the centre line and box boundaries indicate mean ± s.e.m. All *P* values were calculated using a two-tailed Mann–Whitney *U*-test.

neurons in the POA[16–18,46], our results suggest the possibility that avMLPA neurons that express *Vglut2* or *Adcyap1* (avMLPA^{Vglut2} and avMLPA^{Adcyap1} neurons, respectively) regulate the decrease in core body temperature that is associated with natural torpor. To directly test whether stimulation of these neurons is sufficient to phenocopy the decrease in body temperature observed during natural torpor, we used Vglut2-IRES-Cre and Adcyap1-2A-Cre mice and expressed the excitatory Gq-DREADD in avMLPA^{Vglut2} or avMLPA^{Adcyap1} neurons (Fig. 4a). Chemogenetic activation of avMLPA^{Adcyap1} or avMLPA^{Vglut2} neurons resulted in a rapid decrease in core body temperature (5.9 ± 0.4 °C and 4.9 ± 1.0 °C, respectively) and in gross motor activity (58.7 ± 8.4% and 53.3 ± 11.8%, respectively; Fig. 4a–c, Extended Data Fig. 8a–f), which phenocopied the stimulation of torpor-TRAPed avMLPA neurons and effectively recapitulated the decrease in core body temperature and activity observed during natural torpor. By contrast, the chemogenetic activation of avMLPA^{Vgat} neurons in Vgat-IRES-Cre mice led to no significant change in core body temperature (0.2 ± 0.3 °C, $P = 0.48$, Fig. 4a–c, Extended Data Fig. 8b).

Silencing of torpor-associated neurons

Although these findings are consistent with the idea that torpor-active avMLPA^{Adcyap1} and avMLPA^{Vglut2} cells are critical torpor-inducing neurons in the avMLPA, it remained possible that these cells simply constituted a part of the core circuitry that controls homeostatic body temperature, rather than being mediators of torpor entry. To directly address the role of avMLPA^{Adcyap1} and avMLPA^{Vglut2} neuronal activity in basal homeostatic thermoregulation, as well as the natural process of fasting-induced torpor, we targeted the expression of a virally encoded, Cre-dependent tetanus toxin light chain[47] (AAV-Flex-TeLC)—a derivative of a potent neurotoxin that eliminates synaptic transmission—to avMLPA^{Adcyap1} or avMLPA^{Vglut2} neurons (Fig. 4d). By expressing Gq-DREADD and TeLC in the same neurons and stimulating them with CNO, we were able to first verify that TeLC expression effectively inhibited synaptic transmission in these cells, to the extent that it blocked the decrease in core body temperature associated with the chemogenetic stimulation (Extended Data Fig. 8g–i). Experiments with fed mice showed that avMLPA^{Adcyap1} or avMLPA^{Vglut2} silencing had no significant effect on normal homeostatic body temperature, including its circadian rhythm (Fig. 4e, f, Extended Data Fig. 8j–o). By contrast, mice in which avMLPA^{Vglut2} neurons were selectively silenced showed profound disruption of fasting-induced torpor (Fig. 4e). Specifically, the decrease in core body temperature associated with fasting-induced torpor was significantly diminished after the injection of AAV-Flex-TeLC, as compared with that observed in the same mice before injection (pre-fast, $P = 0.010$) or in mice injected with a control AAV vector (control fast, $P = 0.018$, Extended Data Fig. 8j). The kinetics of the torpor-associated decrease in body temperature were also significantly altered by avMLPA^{Vglut2} silencing: mice reached their lowest body temperature after 22 ± 1 h, compared with 14 ± 1 h for control mice (Extended Data Fig. 8k, $P = 9.2 \times 10^{-3}$). Silencing of avMLPA^{Adcyap1} neurons—a subset of avMLPA^{Vglut2} neurons—also altered natural torpor (Fig. 4f, Extended Data Fig. 8l): normal torpor bouts were replaced with a gradual decrease in core body temperature, and an additional 6 h was required to reach a similar degree of hypothermia compared with normal torpor (Extended Data Fig. 8m, $P < 7 \times 10^{-3}$). It must be noted that the injected site in some cases included areas that extended beyond the avMLPA. In addition, using in situ hybridization techniques, we found that only a subset (43 ± 5%) of avMLPA *Adcyap1*$^{+}$ neurons expressed TeLC, perhaps due to inefficient viral transduction or Cre-mediated recombination. This may have led to the incomplete silencing of torpor-regulating avMLPA^{Adcyap1} neurons, and could be a reason for the incomplete elimination of fasting-induced torpor responses that we observed in this paradigm (Extended Data Fig. 8p, q). Nevertheless, these results suggest that the activity of avMLPA^{Adcyap1} neurons is necessary for the normal pattern of torpor thermoregulation—characterized by a rapid decrease in core body temperature, its maintenance and subsequent re-warming—and that, perhaps together with other glutamatergic neurons, avMLPA^{Adcyap1} neurons are required for the depth of hypothermia that is observed during natural torpor.

Dynamics of avMLPA^{Adcyap1} neurons in torpor

Active involvement of glutamatergic *Adcyap1*$^{+}$ avMLPA neurons in the initiation of torpor should be reflected by acute torpor-associated changes in avMLPA^{Adcyap1} neuronal firing patterns. Moreover, the characterization of avMLPA neuronal firing patterns during natural torpor might provide further insight into the underlying functional role of this circuit in torpor regulation. For example, natural avMLPA firing patterns might encode a caloric deficit, circadian time, rate of change in body temperature, time of acute torpor onset or maintenance of torpor. To distinguish between these possibilities, we expressed the calcium reporter[48] GCaMP6s in avMLPA^{Adcyap1} neurons, installed an optical fibre in the region containing these neurons, and continuously monitored neuronal calcium transients over 7–12 h in freely moving fasted mice as they entered torpor (Fig. 4g, Extended Data Fig. 9a, b, Methods). This analysis did not show a gradual change in neuronal activity correlated with caloric restriction or circadian time, or a transient pattern of activity that correlated with only the onset of torpor. Instead, we observed a marked change in neuronal activity that coincided with torpor entry and persisted until the mice began to exit the torpid state (Fig. 4h–l, Extended Data Fig. 9c, d, Methods). This torpor-associated activity pattern was characterized by an 8.0 ± 1.8% decrease in the baseline signal ($P = 3 \times 10^{-3}$) and a 19.0 ± 7.0-fold increase in the frequency of highly prominent Ca^{2+}-dependent peaks ($P = 2 \times 10^{-4}$, Fig. 4m, n, Extended Data Fig. 9d–g) compared with the non-torpor state. Notably, this distinct pattern of avMLPA^{Adcyap1} neuronal activity alone was sufficient to accurately model and determine when mice entered, maintained and exited torpor (Extended Data Fig. 9h–k), which suggests that avMLPA^{Adcyap1} neurons encode information specific to torpor entry and maintenance.

To exclude the possibility that the avMLPA^{Adcyap1} neurons were merely responding to an acute change in body temperature, we recorded their activity in fed mice in which hypothermia had been artificially induced by administration of the adenosine receptor agonist N^6-cyclohexyladenosine (CHA)[49]. Chemically induced hypothermia resulted in a similar decrease in baseline signal (7.2 ± 1.7%) to that seen in torpor, perhaps due to decreased neuronal discharge at lower body temperature. However, chemically induced hypothermia failed to produce the prominent Ca^{2+}-dependent peaks that are associated with fasting-induced torpor, suggesting that the large Ca^{2+} transients observed in avMLPA^{Adcyap1} neurons may encode information that is relevant specifically to torpor and not simply to hypothermia or cooling (Extended Data Fig. 10a–e). Whether this pattern of Ca^{2+} transients reflects the existence of distinct neuronal subpopulations—or perhaps burst firing of a single population of responsive neurons—remains unclear, as does its function in the process of torpor entry.

To investigate whether torpor-active avMLPA^{Adcyap1} neurons are distinct from previously identified populations of warm-sensitive neurons in the POA[16–18,46], we next challenged the mice in which we had previously observed torpor-related avMLPA^{Adcyap1} neuronal activity with either warm (37 °C) or cold (10 °C) environments. Unlike previously described warm-sensitive neurons, avMLPA^{Adcyap1} neurons showed no significant changes in activity in the warm environment; however, they did display sensitivity to a cold environment, suggesting the existence of several functionally distinct subpopulations of POA^{Adcyap1} neurons (Extended Data Fig. 10f–i). In summary, we show that avMLPA^{Adcyap1} neurons both contribute to and are necessary for the natural decrease in body temperature that is observed during torpor, and encode a unique pattern of broad and highly prominent Ca^{2+} transients as mice enter and sustain torpor. This suggests that the activity of these neurons is critical to the natural process of torpor initiation and maintenance in mice.

Discussion

Our study examines the mechanisms underlying a complex naturalistic behaviour by using several recent technological advances, including the FosTRAP approach, machine learning-enabled image registration, snRNA-seq and long-term recordings of neuronal activity. We implicate specific neuronal cell types within a defined brain region in the regulation of torpor, one of the most extreme and poorly understood physiological adaptations in homeothermic animals. We discover that neural activity alone is sufficient to induce several key features of torpor, including decreased locomotion and a profoundly lowered metabolic rate and core body temperature in mice that are not calorically restricted. We identify the avMLPA of the hypothalamus as a torpor control centre, and *Vglut2*$^{+}$*Adcyap1*$^{+}$ neurons as a central torpor-regulating neuronal population. We observe that the activity of these neurons changes markedly when mice naturally enter torpor, and that this both contributes to and is necessary for the precipitous decrease in core body temperature that is observed during natural torpor. How these torpor-regulating neurons integrate information about internal states and environmental experience remains to be understood. Glutamatergic avMLPA^{Adcyap1} neurons receive information about decreasing ambient temperatures, leading us to speculate that these neurons might integrate information about environmental conditions with information about the internal energy reserves of the animal, so as to control when the animal enters torpor. Transcriptome-wide analysis of *Vglut2*$^+$*Adcyap1*$^+$ neurons identified expression of the leptin receptor in two of the *Vglut2*$^+$*Adcyap1*$^+$ neuronal subtypes, suggesting the possibility that circulating leptin levels might modulate the activity of torpor-regulating neurons and provide a potential mechanism for how information about decreased energy reserves in fasted mice could be conveyed to torpor-regulating avMLPA$^{Vglut2/Adcyap1}$ neurons[27,30] (Extended Data Fig. 5j). In addition, given that entry into torpor requires hours of fasting, and involves many neuronal populations in which the FOS transcription factor is induced, one possibility is that activity-dependent gene transcription has a mechanistic role in establishing a state that is permissive for torpor entry.

Because torpor is a dynamic and complex behaviour that involves profound changes in body temperature, metabolic rate, locomotion, perception, breathing and heart rate[7,10], the engagement of other neuronal populations and brain regions—including additional torpor-associated neurons identified here by the FosTRAP approach and by snRNA-seq—is likely to be involved in orchestrating the full program of natural torpor entry, maintenance and arousal. The elucidation of torpor-regulating neuronal circuitry will enable an investigation of the mechanisms by which this circuit inhibits, circumvents or inverts normal cold defensive thermoregulatory processes[1,2,50]. This should provide insight into why only certain mammalian species have the ability to enter torpor, and whether such a hypometabolic state could be induced in species that typically do not enter torpor. For example, although rats and humans do not naturally enter torpor, a study has recapitulated attributes of torpor in rats[39], suggesting the possibility of inducing a similar state in humans. Future investigations should deliver further advances through the study and manipulation of these ancient adaptations of homeotherm biology.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2387-5.

1. Tan, C. L. & Knight, Z. A. Regulation of body temperature by the nervous system. *Neuron* **98**, 31–48 (2018).
2. Morrison, S. F. & Nakamura, K. Central mechanisms for thermoregulation. *Annu. Rev. Physiol*. **81**, 285–308 (2019).
3. Heller, H. C. & Hammel, H. T. CNS control of body temperature during hibernation. *Comp. Biochem. Physiol. A* **41**, 349–359 (1972).
4. Sunagawa, G. A. & Takahashi, M. Hypometabolism during daily torpor in mice is dominated by reduction in the sensitivity of the thermoregulatory system. *Sci. Rep*. **6**, 37011 (2016).
5. Geiser, F. Metabolic rate and body temperature reduction during hibernation and daily torpor. *Annu. Rev. Physiol*. **66**, 239–274 (2004).
6. Morton, S. R. Torpor and nest-sharing in free-living S *minthopsis crassicaudata* (Marsupialia) and *Mus musculus* (Rodentia). *J. Mamm*. **59**, 569–575 (1978).
7. Heldmaier, G., Ortmann, S. & Elvert, R. Natural hypometabolism during hibernation and daily torpor in mammals. *Respir. Physiol. Neurobiol*. **141**, 317–329 (2004).
8. Hoffstaetter, L. J., Bagriantsev, S. N. & Gracheva, E. O. TRPs et al.: a molecular toolkit for thermosensory adaptations. *Pflugers Arch*. **470**, 745–759 (2018).
9. Barnes, B. M. Freeze avoidance in a mammal: body temperatures below 0 °C in an Arctic hibernator. *Science* **244**, 1593–1595 (1989).
10. Swoap, S. J. & Gutilla, M. J. Cardiovascular changes during daily torpor in the laboratory mouse. *Am. J. Physiol. Regul. Integr. Comp. Physiol*. **297**, R769–R774 (2009).
11. Revel, F. G. et al. The circadian clock stops ticking during deep hibernation in the European hamster. *Proc. Natl Acad. Sci. USA* **104**, 13816–13820 (2007).
12. Teague, R. S. & Ranson, S. W. The role of the anterior hypothalamus in temperature regulation. *Am. J. Physiol*. **117**, 562–570 (1936).
13. Zaretskaia, M. V., Zaretsky, D. V., Shekhar, A. & DiMicco, J. A. Chemical stimulation of the dorsomedial hypothalamus evokes non-shivering thermogenesis in anesthetized rats. *Brain Res*. **928**, 113–125 (2002).
14. Schneeberger, M. et al. Regulation of energy expenditure by brainstem GABA neurons. *Cell* **178**, 672–685 (2019).
15. Boulant, J. A. & Hardy, J. D. The effect of spinal and skin temperatures on the firing rate and thermosensitivity of preoptic neurones. *J. Physiol. (Lond.)* **240**, 639–660 (1974).
16. Tan, C. L. et al. Warm-sensitive neurons that control body temperature. *Cell* **167**, 47–59 (2016).
17. Yu, S. et al. Glutamatergic preoptic area neurons that express leptin receptors drive temperature-dependent body weight homeostasis. *J. Neurosci*. **36**, 5034–5046 (2016).
18. Harding, E. C. et al. A neuronal hub binding sleep initiation and body cooling in response to a warm external stimulus. *Curr. Biol*. **28**, 2263–2273 (2018).
19. Nakayama, T., Eisenman, J. S. & Hardy, J. D. Single unit activity of anterior hypothalamus during local heating. *Science* **134**, 560–561 (1961).
20. Wang, T. A. et al. Thermoregulation via temperature-dependent PGD_2 production in mouse preoptic area. *Neuron* **103**, 309–322 (2019).
21. Feketa, V. V., Nikolaev, Y. A., Merriman, D. K., Bagriantsev, S. N. & Gracheva, E. O. CNGA3 acts as a cold sensor in hypothalamic neurons. *eLife* **9**, e55370 (2020).
22. Herrington, L. P. The heat regulation of small laboratory animals at various environmental temperatures. *Am. J. Physiol*. **129**, 123–139 (1940).
23. Reitman, M. L. Of mice and men – environmental temperature, body temperature, and treatment of obesity. *FEBS Lett*. **592**, 2098–2107 (2018).
24. Padilla, S. L. et al. Agouti-related peptide neural circuits mediate adaptive behaviors in the starved state. *Nat. Neurosci*. **19**, 734–741 (2016).
25. Jensen, T. L., Kiersgaard, M. K., Sørensen, D. B. & Mikkelsen, L. F. Fasting of mice: a review. *Lab. Anim*. **47**, 225–240 (2013).
26. Hitrec, T. et al. Neural control of fasting-induced torpor in mice. *Sci. Rep*. **9**, 15462 (2019).
27. Gavrilova, O. et al. Torpor in mice is induced by both leptin-dependent and -independent mechanisms. *Proc. Natl Acad. Sci. USA* **96**, 14623–14628 (1999).
28. van der Vinne, V., Bingaman, M. J., Weaver, D. R. & Swoap, S. J. Clocks and meals keep mice from being cool. *J. Exp. Biol*. **221**, jeb179812 (2018).
29. Himms-Hagen, J. Food restriction increases torpor and improves brown adipose tissue thermogenesis in ob/ob mice. *Am. J. Physiol. Endocrinol. Metab*. **248**, E531–E539 (1985).
30. Swoap, S. J., Gutilla, M. J., Liles, L. C., Smith, R. O. & Weinshenker, D. The full expression of fasting-induced torpor requires β3-adrenergic receptor signaling. *J. Neurosci*. **26**, 241–245 (2006).
31. Greenberg, M. E., Ziff, E. B. & Greene, L. A. Stimulation of neuronal acetylcholine receptors induces rapid gene transcription. *Science* **234**, 80–83 (1986).
32. *Allen Mouse Brain Atlas* (Allen Institute for Brain Science, 2019); http://mouse.brain-map.org
33. Wu, Q. et al. The temporal pattern of cfos activation in hypothalamic, cortical, and brainstem nuclei in response to fasting and refeeding in male mice. *Endocrinology* **155**, 840–853 (2014).
34. Andermann, M. L. & Lowell, B. B. Toward a wiring diagram understanding of appetite control. *Neuron* **95**, 757–778 (2017).
35. Nectow, A. R. et al. Identification of a brainstem circuit controlling feeding. *Cell* **170**, 429–442 (2017).
36. Allen, W. E. et al. Thirst-associated preoptic neurons encode an aversive motivational drive. *Science* **357**, 1149–1155 (2017).
37. Zhu, H. et al. Cre-dependent DREADD (Designer Receptors Exclusively Activated by Designer Drugs) mice. *Genesis* **54**, 439–446 (2016).
38. Geiser, F., Currie, S. E., O'Shea, K. A. & Hiebert, S. M. Torpor and hypothermia: reversed hysteresis of metabolic rate and body temperature. *Am. J. Physiol. Regul. Integr. Comp. Physiol*. **307**, R1324–R1329 (2014).
39. Tupone, D., Madden, C. J. & Morrison, S. F. Central activation of the A1 adenosine receptor (A1AR) induces a hypothermic, torpor-like state in the rat. *J. Neurosci*. **33**, 14512–14525 (2013).
40. McKinley, M. J. et al. The median preoptic nucleus: front and centre for the regulation of body fluid, sodium, temperature, sleep and cardiovascular homeostasis. *Acta Physiol. (Oxf.)* **214**, 8–32 (2015).
41. Moffitt, J. R. et al. Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science* **362**, eaau5324 (2018).
42. Sherin, J. E., Shiromani, P. J., McCarley, R. W. & Saper, C. B. Activation of ventrolateral preoptic neurons during sleep. *Science* **271**, 216–219 (1996).
43. Chung, S. et al. Identification of preoptic sleep neurons using retrograde labelling and gene profiling. *Nature* **545**, 477–481 (2017).
44. Li, S.-B. & de Lecea, L. The hypocretin (orexin) system: from a neural circuitry perspective. *Neuropharmacology* **167**, 107993 (2020).
45. Habib, N. et al. Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat. Methods* **14**, 955–958 (2017).
46. Zhao, Z.-D. et al. A hypothalamic circuit that controls body temperature. *Proc. Natl Acad. Sci. USA* **114**, 2042–2047 (2017).
47. Yamamoto, M. et al. Reversible suppression of glutamatergic neurotransmission of cerebellar granule cells in vivo by genetically manipulated expression of tetanus neurotoxin light chain. *J. Neurosci*. **23**, 6759–6767 (2003).
48. Chen, T.-W. et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).
49. Carlin, J. L. et al. Hypothermia in mouse is caused by adenosine A_1 and A_3 receptor agonists and AMP via three distinct mechanisms. *Neuropharmacology* **114**, 101–113 (2017).
50. Tupone, D., Cano, G. & Morrison, S. F. Thermoregulatory inversion: a novel thermoregulatory paradigm. *Am. J. Physiol. Regul. Integr. Comp. Physiol*. **312**, R779–R786 (2017).

Methods

Mice

Animal experiments were approved by the National Institutes of Health and Harvard Medical School Institutional Animal Care and Use Committee, following ethical guidelines described in the US National Institutes of Health Guide for the Care and Use of Laboratory Animals. For initial torpor experiments, we used adult (6–10-week-old) C57BL/6J mice (The Jackson Laboratory, Stock 000664). To generate FosTRAP-Gq mice we crossed Fos2A-iCreER (TRAP2) mice (The Jackson Laboratory, Stock 030323) with R26-LSL-Gq-DREADD mice (The Jackson Laboratory, Stock 026220) and used adult (6–18-week-old) male and female F_1 progeny. For viral injections we used Fos2A-iCreER (TRAP2) mice (The Jackson Laboratory, Stock 030323), Adcyap1-2A-Cre mice (The Jackson Laboratory, Stock 030155), Vglut2-ires-Cre mice (The Jackson Laboratory, Stock 028863) and Vgat-ires-Cre mice (The Jackson Laboratory, Stock 028862). All mice were housed at 22 °C under a standard 12 h light/dark cycle.

No statistical methods were used to predetermine the sample size. Mice were randomly assigned to experimental groups before surgery. Where possible, investigators were blinded during analysis.

Telemetric monitoring of core body temperature and gross motor activity

Mice with genotypes of interest were singly housed and implanted abdominally with telemetric temperature and activity probes (Starr Life Science VV-EMITT-G2). After at least four days of recovery, mice were recorded in standard cages placed onto a radiofrequency receiver platform (Starr Life Science ER4000). Core body temperature and gross motor activity were logged every 60 s.

Measurements of oxygen consumption

Oxygen consumption was measured using the Columbus Instruments Comprehensive Lab Animal Monitoring System (CLAMS) at the Brigham and Women's Hospital or the Beth Israel Deaconess Medical Center Metabolic Core. Individual mice were placed into metabolic cages enabling the measurement of oxygen consumption (recorded as VO_2, the volume of O_2 consumed per unit time). The CLAMS system is compatible with simultaneous measurements of core body temperature and gross motor activity using the implanted telemetric probes.

Torpor induction

Adult (6–18-week-old) mice were singly housed before the induction of torpor. Each mouse was moved to a new individual cage containing water and nesting material but devoid of bedding and food at the beginning of the dark cycle. Initial bouts of torpor were observed after approximately 8 h of fasting. Mice were returned to their standard cages containing food 24 h after the start of the fast. The ambient temperature of the facility was maintained at around 22 °C.

TRAPing

To label neurons that are active during torpor, we used mice harbouring a tamoxifen-dependent form of Cre recombinase (CreERT2) driven from the *Fos* locus ($Fos^{2A\text{-}iCreERT2}$, FosTRAP[36]). When these mice are fasted to enter torpor, the neurons that are active induce FOS and CreERT2. When these mice are exposed to 4-OHT (Sigma-Aldrich, H6278), the CreERT2 will translocate to the nucleus. Once translocated, CreERT2 can recombine the genomically encoded Cre-dependent R26-LSL-Gq-DREADD allele or the virally introduced Cre-dependent AAV-DIO-Gq-DREADD–mCherry construct selectively in the FOS-expressing cells. This recombination leads to the persistent expression of Gq-DREADD and allows for the activation of these specific 'TRAPed' neurons at a later time by the administration of CNO[51].

4-OHT solution was prepared by initially dissolving 10 mg of 4-OHT in 500 µl of 100% ethanol, then adding 450 µl of a 1:4 mixture of castor oil:sunflower oil and vortexing. The ethanol was removed via vacuum centrifugation, and the remaining 4-OHT solution was diluted with the same 1:4 mixture of castor oil:sunflower oil to a final concentration of approximately 6.25 mg ml^{-1}. For TRAPing, each mouse was injected intraperitoneally with 50 mg kg^{-1} of 4-OHT during torpor entry.

CNO solution was prepared by initially dissolving CNO hydrochloride (Sigma-Aldrich, SML2304) in H_2O to a stock solution of 100 mM. The stock solution was diluted with PBS to a final concentration of 0.6 mM, and approximately 250 µl was injected intraperitoneally per mouse for a final injection concentration of 2 mg kg^{-1}.

Immunofluorescent staining

Mice were euthanized by transcardial perfusion of 10 ml PBS followed by 10 ml of 4% paraformaldehyde (PFA). Brains were extracted, post-fixed overnight with 4% PFA at 4 °C and then embedded in PBS with 3% agarose. Brains were sliced on a vibratome (Leica VT1000S) into 50-µm coronal sections. Coronal sections were washed three times with PBS containing 0.3% TritonX-100 (PBST) and blocked for 1 h at room temperature with PBST containing 5% donkey serum (blocking buffer). Sections were incubated overnight at 4 °C with primary antibodies diluted in blocking buffer, washed again three times with PBST, and incubated for 1 h at room temperature with secondary antibodies diluted in blocking buffer. After washing twice in PBST and once in PBS, samples were mounted onto SuperFrost Plus glass slides (VWR) using DAPI Fluoromount-G.

Detection of FOS$^+$ and Gq-DREADD-HA$^+$ cells

Tissues were processed as indicated in the section 'Immunofluorescent staining' using the following reagents: primary antibodies: rabbit anti-FOS antibody 1:2,000 (Cedarlane, 226003(SY)) and rabbit anti-haemagglutinin (HA) antibody 1:1,000 (Cell Signaling Technology, 3724S). Secondary antibody: 1:500 donkey anti-rabbit 647 secondary antibody (Life Technologies, A31573).

Sections were imaged on an Olympus BX61VS microscope using a UPlanSApo 10 × 0.4 objective (Harvard NeuroDiscovery Center).

Image registration to the Allen Brain Atlas and analysis

Imaged immunofluorescent brain slices were converted to TIF format and organized in sequential anterior to posterior order. To quantify the average Gq-DREADD-HA signal and count FOS$^+$ cells over different brain regions, the brain slices were registered with the Allen Brain Atlas[32] as described previously[52–55], with some custom modifications. There are three main parts to the pipeline: pre-processing, stack registration and plane alignment. The pre-processing stage includes downsizing image files, using a previously trained machine-learning algorithm to compute the brain slice borders, generating a mask and edge map, and detecting puncta if necessary. The stack-registration stage aligns the dataset with itself using the vertical line of symmetry as a guide. In the plane-assignment stage, each internally registered brain slice is paired to the Allen Brain Atlas using an estimate of the bregma value and a method similar to sequence alignment using dynamic programming. After pairwise registration to the Allen Brain Atlas, experimental brain slices are further adjusted using local nonlinear deformations to maximize the fit between the experimental image and the image from the Allen Brain Atlas. Then, the signal intensity and the number of spots are quantified for each brain region defined by the Allen Brain Atlas. Finally, the volume of each brain region is calculated so that the signal density (signal intensity divided by volume) and density of spots (number of spots divided by volume) could be determined.

Several parameters in the pipeline were adjusted to optimally process our experimental brain slices. For both FOS and Gq-DREADD-HA signal quantification, a boundary erosion with radius 1 (after downsizing) was implemented to exclude quantification at the edge of the brain slices, where inaccuracies during the experimental preparation may occur. Estimated bregma values for the most anterior and posterior brain slices were also adjusted on the basis of each sample. To detect individual FOS$^+$ nuclei via Laplacian of Gaussian filtering of the image, we found and distinguished the local maxima that are FOS puncta from local maxima in the background. The distribution of local maxima in the background was computed, and the threshold distance to background distribution was set to 7 to indicate that only spots 7 standard deviations away from the mean of the background distribution were selected. For selected spots (putative FOS$^+$ nuclei), their correlation with the ideal circular spot was measured, and any spot with a correlation below 0.5 was eliminated. These parameters were ignored when the FosTRAP-Gq-HA brain signal was quantified, as the haemagglutinin signal was not nuclear and was instead distributed across the cell body and neuronal processes, and thus could not be modelled and quantified as a circular spot.

Viral constructs

AAV8-hSyn-DIO-Gq–mCherry (Addgene, 44361-AAV8) and AAV1-Syn-Flex-GCaMP6s-WPRE-SV40 (Addgene, 100845-AAV1) were obtained from Addgene. AAV2/1-hSyn-Flex-TeLC-eYFP[56] was prepared through Boston Children's Hospital Viral Core. All viruses were diluted with PBS to a final concentration between 5×10^{12} and 1×10^{13} genome copies per ml before stereotaxic delivery into the mouse brain.

Stereotactic viral injection and photometry fibre implantation

For injections, mice were anaesthetized with 3% isoflurane and placed in a stereotaxic head frame (Kopf Instrument, model 1900). Coordinates AP +0.4 mm, ML ±0.5 mm, DV −5.1 mm relative to bregma, were used for all avMLPA injection. Unless otherwise specified, all experiments were carried out with bilateral injections. An air-based injection system built with Digital Manometer (Grainger, 9LHH8) was used to infuse the virus. The virus was infused at approximately 100 nl min^{-1}, and the needle was kept at the injection site for 10 min before withdrawing. For chemogenetic stimulation, about 25–100 nl of AAV8-hSyn-DIO-Gq–mCherry was delivered into the region of interest. For fibre photometry recordings, 200 nl of a mixture (1:1 ratio) of AAV8-hSyn-DIO-Gq–mCherry and AAV1-Syn-Flex-GCaMP6s-WPRE-SV40 was delivered into the region of interest, either unilaterally or bilaterally. For fibre photometry recording, Mono Fibre-optic cannulas (Doric, MFC_200/230-0.37_###_MF1.25_FLT) were implanted 200 µm above the injection site. The fibre was fixed to the skull with Loctite 454 Instant Adhesive and further covered by dental cement to ensure the stability of the implant.

Mapping of virally injected hypothalamic nuclei and correlation with chemogenetically induced changes in core body temperature

Viral injection, sample preparation and imaging. To minimize initial differences in body weight between mice, which might influence the effects on core body temperature and confound our screen across brain regions, we used a cohort of fifty-four 6–10-week-old female FosTRAP mice. Mice were bilaterally injected with AAV8-hSyn-DIO-Gq–mCherry, allowed to recover for 4–10 days and then fasted and injected with 4-OHT during torpor to TRAP active neurons. Subsequently (3–14 days later), mice were administered CNO to stimulate fast-TRAPed neurons within the virally injected regions and their core body temperature and activity were recorded. Given that different mice were injected in different hypothalamic areas and that no two injections will be identical due to the subtle variabilities in the injection site and viral spread, all mice were euthanized to map post hoc the exact brain regions that were transduced and contained TRAPed neurons. Mice were euthanized by transcardial perfusion of 10 ml PBS followed by 10 ml of 4% PFA. Brains were extracted, post-fixed overnight with 4% PFA at 4 °C and then embedded in PBS with 3% agarose. Brains were sliced on a vibratome (Leica VT1000S) into 50 µm coronal sections, mounted onto glass slides and imaged on an Olympus BX61VS microscope using a UPlanSApo 10 × 0.4 objective (Harvard NeuroDiscovery Center). The expression of AAV-DIO-Gq-DREADD–mCherry was detected by its endogenous fluorescence.

Image analysis. Viral mCherry expression was quantified across 277 hypothalamic regions, noting semiquantitatively for each region and each hemisphere whether the viral expression was 0 (none), 1 (minimal), 2 (partial) or 3 (total; Fig. 2b, Supplementary Table 3). The analysis was performed blinded to the effect on core body temperature that was previously observed in each mouse.

Correlation with changes in core body temperature. To assign a single numeric value for each hypothalamic region in each mouse, we added the semiquantitative transduction values from the two hemispheres. For correlation analysis, we analysed 226 of the 277 brain regions that were transduced in at least three mice. For each region independently, we determined the Pearson correlation across all mice between the viral expression in that region and the decrease in core body temperature that was observed. FDR-corrected *q* values are shown in Extended Data Fig. 2a–c. For each region independently, we also calculated the average (across mice) minimum temperature that was observed after the chemogenetic stimulation of all mice in which the region was hit (viral expression > 0) and all mice in which the regions were missed (viral expression = 0). Values are plotted in Extended Data Fig. 2a–c.

Nuclear isolation for snRNA-seq

Mice were anaesthetized with 3% isoflurane and transcardially perfused with cold choline dissection media. The brains were extracted and sectioned into 300-µm coronal sections. Sections containing the MPA were microdissected to isolate the avMLPA.

Single-nuclei suspensions for droplet-based snRNA-seq[57–59] were generated as described previously[60], with minor modifications. The avMLPA was dissected and placed into a Dounce with homogenization buffer (0.25 M sucrose, 25 mM KCl, 5 mM $MgCl_2$, 20 mM Tricine-KOH, pH 7.8, 1 mM DTT, 0.15 mM spermine, 0.5 mM spermidine, protease inhibitors). The sample was homogenized using a tight pestle with 10 strokes. IGEPAL solution (5%, Sigma) was added to a final concentration of 0.32%, and 5 additional strokes were performed. The homogenate was filtered through a 40-µm filter, and OptiPrep (Sigma) was added to a final concentration of 25% iodixanol. The sample was layered onto an iodixanol gradient and centrifuged at 10,000*g* for 18 min as previously described[59,60]. Nuclei were collected between the 30% and 40% iodixanol layers and diluted to 80,000-100,000 nuclei per ml for encapsulation. All buffers contained 0.15% RNasin Plus RNase Inhibitor (Promega) and 0.04% BSA.

avMLPA snRNA-seq

snRNA-seq library preparation and sequencing. Single nuclei were captured and barcoded for whole-transcriptome libraries using the 10X Genomics Chromium v2 platform according to the manufacturer's recommendations, collecting one library of approximately 10,000 nuclei from each of the 5 mice. In brief, single nuclei along with single primer-carrying hydrogels were captured into droplets using a microfluidic device. Each hydrogel carried oligodT primers with a unique cell barcode. Nuclei were lysed and the cell-barcode-containing primers released from the hydrogel, initiating reverse transcription and barcoding of all cDNA in each droplet. Next, the emulsions were broken and cDNA across approximately 10,000 nuclei were pooled into the same library. The cDNA was amplified, fragmented and adapters were added for sequencing on a Nextseq 500 benchtop DNA sequencer (Illumina).

For enrichment of virally derived transcripts, a fraction (1 µl) of the non-fragmented cDNA was PCR-amplified. The forward primer (5′-GCATGGACGAGCTGTACA) was designed to anneal to the sequence uniquely present at the 3′ terminus of mCherry. The reverse primer (5′-CTACACGACGCTCTTCCG) was designed to anneal to the R1 sequence, uniquely present at the 5′ terminus, to the 10X barcode and unique molecular identifier (UMI) sequence that were introduced during the reverse transcription. The result of the PCR is preferential amplification of the viral-derived transcripts, while simultaneously retaining the cell-barcode sequence necessary to assign each viral transcript to a particular cell or nucleus. After PCR amplification (18 cycles, Hot Start High-Fidelity Q5 polymerase, NEB, M0494S), the product was purified using 0.6X SPRI select reagent (Thermo Fisher Scientific, NC0406406) and 1 µl of the 50 µl eluent was used in a second PCR reaction. The forward primer for the second PCR (5′-GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTgtaaggcgcgccataac) was designed to anneal to the sequence uniquely present between the *mCherry* and *loxP* sites of the AAV-DIO-Gq-DREADD–mCherry vector. In addition, this primer introduced the R2 sequence necessary for later library amplification. The reverse primer (5′-CTACACGACGCTCTTCCG) was the same as in the first PCR reaction. The result of the PCR is again preferential amplification of the viral-derived transcripts, while simultaneously retaining the cell-barcode sequence necessary to assign each viral transcript to a particular cell or nucleus. After PCR amplification (18 cycles, Hot Start High-Fidelity Q5 polymerase, NEB, M0494S), all the libraries were indexed according to the 10X protocol, pooled, and sequenced on a Nextseq 500 benchtop DNA sequencer (Illumina).

snRNA-seq sample mapping and viral barcode deconvolution by cell. The 10X Genomics package cellranger 3.1.0 was used to map transcripts to the mm10 reference mouse genome. Because snRNA-seq captures unspliced pre-mRNA as well as mature mRNA, we created a custom ‘pre-mRNA’ gene annotation file, in which all features are recoded as exons, enabling us to map unspliced reads to these artificial exons and to assign them to the corresponding genes. Feature barcoding with custom features was used to assign viral mRNA reads to cells and distinguish between the transcript sequence from the initial AAV-DIO-Gq-DREADD–mCherry vector and the transcript sequence produced by the Cre-recombined vector in TRAPed neurons.

Doublet removal, embedding and identification of main cell classes. Microfluidic encapsulation of nuclei results in some droplets containing more than one nucleus. During reverse transcription, transcripts from co-encapsulated nuclei would be labelled with the same cell barcode, effectively creating a hybrid cell or a ‘doublet’. To detect and remove predicted doublets using bioinformatics, each of the five libraries were independently processed using the *scrublet* package[61] with the following commands:

scrub = scr.Scrublet(counts_matrix, expected_doublet_rate = 0.10)

doublet_scores, predicted_doublets = scrub.scrub_doublets(min_counts = 2, min_cells = 3, min_gene_variability_pctl = 85, n_prin_comps = 30).

Doublet-removal eliminated 2,037 cells, leaving 44,669 nuclei across five libraries for further processing. We recovered an average of 2,083 unique non-viral transcripts per nucleus, representing 1,286 unique genes (Extended Data Fig. 3). Data from all nuclei were analysed simultaneously and virally derived sequences were removed for the purposes of embedding, clustering and cell-type identification. The R software package Seurat 3.1[58,62] was used to assign cells to clusters. Any genes that were expressed in fewer than three cells were removed from the analysis. Data was normalized with the number of transcripts per cell and the per cent transcripts derived from the mitochondrial genes regressed using the SCTransform function with default parameters. Initially, the 3,000 most variable genes were identified.

Tissue dissociation has been shown to increase the expression of mitochondrial RNAs, ribosomal-protein RNAs, and immediate–early genes[63–65]. To minimize the influence of these tissue-processing-induced genes on the identification and molecular characterization of avMLPA cells, we removed from the initial list of 3,000 most variable genes any genes that were identified as mitochondrial, ribosomal-protein-encoding or were the immediate–early genes *Fos*, *Fosb*, *Fosl1*, *Fosl2*, *Egr1* and *Npas4*. To further diminish any potential effects of tissue-dissociation-induced expression changes on clustering, we identified and removed from the list of variable genes any genes for which their expression across all cells correlated with any mitochondrial, ribosomal-protein-encoding or immediate–early genes listed above[41,58,66] (Pearson correlation coefficient greater than 0.2 or smaller than −0.2). The final list contained 2,827 variable genes. Next, principal component analysis (PCA) was carried out using the *RunPCA()* function. The *FindNeighbours()* function, using the top 30 principal components (PCs) and the *FindClusters()*, was used to identify the initial 34 clusters. Clusters that were disproportionately derived from a single sample were removed, leaving 39,562 cells. On the basis of the expression of known marker genes, we merged clusters that represented the same cell type. Our final list of cell types was: glutamatergic neurons (Glut), GABAergic neurons (GABA), cholinergic neurons (Chol), astrocytes (Astro), endothelial cells (Endo), microglia (Micro), oligodendrocytes (Oligo) and oligodendrocyte precursor cells (OPCs) (Fig. 3b, c).

Identification of neuronal subtypes. Cells classified as neurons (n = 28,103 cells) were additionally processed to identify neuronal subtypes in the avMLPA. The same Seurat 3.1.0 pipeline was used as described in the section ‘Doublet removal, embedding and identification of main cell classes’. Clustering identified 24 GABAergic, 8 glutamatergic, 3 hybrid and 1 cholinergic neuronal cell type (Fig. 3d). The function *FindAllMarkers(seurat_mat, only.pos = F, min.pct = 0.1, thresh.use = 0.25)* uses a Bonferroni-corrected (across 23,967 genes in the dataset) two-tailed Mann–Whitney U-test to perform differential gene expression analysis and identify markers of each cell type. The top five markers based on fold enrichment are plotted in Extended Data Fig. 4. The top 20 markers are displayed in Supplementary Table 4. Cross-referencing these markers with previously described markers for cell types in the MPA[41] led to the annotation of the 36 neuronal cell types indicated in Fig. 3e.

Hierarchical tree construction. Neuronal cell types were clustered on the basis of average gene expression across all 2,954 variable genes. A distance matrix was then calculated in Euclidean space and hierarchical clustering was carried out using the function *hclust* and the *ward.D* method.

Identification of TRAPed neurons in snRNA-seq data. Feature barcoding with custom features identified for each nucleus the number of viral mRNA reads that were derived from the non-recombined AAV-DIO-Gq-DREADD–mCherry vector and the number of viral mRNA reads from the Cre-recombined AAV-DIO-Gq-DREADD–mCherry vector (Extended Data Fig. 5a). Neuronal nuclei or cells containing three or more virally derived transcripts (17,424) were considered transduced by the AAV. Among them, 343 neuronal nuclei or cells contained three or more virally derived transcripts for which the mRNA sequence indicated that the vector had been recombined by Cre, suggesting that these cells were previously TRAPed and now express the Gq-DREADD–mCherry protein.

Differential gene expression between TRAPed and non-TRAPed *Adcyap1*$^+$ neurons. Differential gene expression analysis was carried out between TRAPed (n = 139 cells) and non-TRAPed (n = 5,848 cells) *Adcyap1*+ neurons in the snRNA-seq dataset. The FindMarkers(PACAP.cells, ident.1 = “TRAP”, ident.2 = “Non-TRAP”, verbose = FALSE) function in the R software package Seurat 3.1 was used to perform a Bonferroni-corrected (across 23,967 genes in the data set) two-tailed Mann–Whitney U-test and identify differentially expressed genes.

Combined fluorescence in situ hybridization and immunofluorescence

Fluorescence in situ hybridization. Mice were euthanized by transcardial perfusion of 10 ml PBS followed by 10 ml of 4% PFA. Brains were extracted, post-fixed overnight with 4% PFA at 4 °C and then incubated in PBS with 30% sucrose for 48 h at 4 °C for cryoprotection. Brains were embedded in tissue freezing medium and frozen in 2-methylbutane that had been cooled by liquid nitrogen. Brains were sliced on a cryostat (Leica CM1950) into 20-µm sections, adhered to SuperFrost Plus slides (VWR) and immediately stored at −80 °C until use. Samples were processed according to the ACD RNAscope Fluorescent Multiplex Assay manual with the following modifications: 500 ml of 1x Antigen retrieval solution was heated to 99–100 °C and maintained at a uniform boil. The slides of fixed frozen brain slices, stored at −80 °C, were immediately placed into a slide rack and slowly submerged into the boiling 1x Antigen retrieval solution for 5 min. Immediately afterwards, the slides were washed 3–5 times by moving the slide rack up and down in Milli-Q H_2O and washed again in 100% EtOH at room temperature. A hydrophobic barrier was drawn around each slice and it was allowed to dry for 1 min at room temperature. From this point on, the procedure followed the standard in situ hybridization protocol for RNAscope Fluorescent Multiplex Assay.

Immunofluorescence. Slides were washed twice for 2 min in PBST (PBS + 0.01% Tween) and blocked by incubation with 1% BSA and 10% donkey serum in PBST for 30 min room temperature. They were then incubated with primary antibody diluted in PBS + 1% BSA overnight at 4 °C and washed three times for 5 min with PBST. Slides were subsequently incubated with secondary antibody diluted in PBS + 1% BSA for 1 h at room temperature, washed with PBST three times for 2 min at room temperature, stained with DAPI (RNAscope), and finally mounted with ProLong Gold antifade reagent.

Determining markers of torpor-TRAPed avMLPA neurons

Brains were processed as indicated in the section 'Combined fluorescence in situ hybridization and immunofluorescence' using the following reagents: for the detection of mCherry$^+$ cells, 1:300 rabbit anti-mCherry (Abcam, ab167453) was used as the primary antibody and 1:500 donkey anti-rabbit 568 (Life Technologies, AB_2534017) as the secondary antibody.

Sample imaging. Sections containing avMLPA were imaged on a Leica SPE confocal microscope using an ACS APO 20×/0.60 IMM CORR objective (Harvard NeuroDiscovery Center). Tiled MPA areas were imaged with a single optical section to avoid counting the same cell across multiple optical sections. Channels were imaged sequentially to avoid any optical crosstalk.

Image analysis. To determine the fraction of TRAPed mCherry$^+$ cells that express each marker gene, in each image mCherry$^+$ cells were manually marked while staying blinded to the in situ hybridization signals (*Adcyap1*, *Vglut2*, *Vgat*). Because mCherry was fused to the Gq-DREADD and thus largely membrane-bound, we reasoned that manually marking mCherry$^+$ cells would provide a more accurate measurement compared to semi-automated algorithms that are optimized for a more focal (nuclear or cytoplasmic) signal. After the identification of mCherry$^+$ cells, for each cell we evaluated whether it appeared positive for markers detected by in situ hybridization, staying blinded to the identity of the marker that was being evaluated. To determine the fraction of marker$^+$ cells that are mCherry$^+$, we additionally counted the total number of marker$^+$ cells.

Mapping anterograde projections of torpor-TRAPed avMLPA neurons

FosTRAP mice that were injected with AAV8-hSyn-DIO-Gq–mCherry and TRAPed during natural torpor were euthanized 2 months later. Brains were processed as indicated in the section 'Immunofluorescent staining' using the following reagents: primary antibody: 1:300 rabbit anti-mCherry antibody (Abcam, ab167453); secondary antibody: donkey anti-rabbit 647 secondary antibody (Life Technologies, A31573). Sections were imaged on an Olympus BX61VS microscope using a UPlanSApo 10× 0.4 objective (Harvard NeuroDiscovery Center).

Silencing avMLPA neurons

The avMLPA of male and female Adcyap1-2A-Cre and Vglut2-IRES-Cre mice was injected with approximately 100–150 nl AAV-DIO-Gq-DREADD–mCherry and AAV-Flex-TeLC–eYFP (1:1 ratio). Control mice were injected with AAV-DIO-Gq-DREADD–mCherry and/or AAV-Flex-GCaMP6s. After 5–14 days of recovery from surgery, mice were fasted to induce torpor and their core body temperature was monitored. The minimum core body temperature is shown in Fig. 4d–f, Extended Data Fig. 8j–m. To analyse the kinetics of torpor, we calculated the time it took each mouse to reach minimum body temperature, requiring that the minimum body temperature is at least as low as the 10th percentile of the body temperature observed in control fed mice (33.2 °C). To avoid counting smaller oscillations in body temperature in this analysis, we classified the mice as not having entered torpor if the body temperature of a fasted mouse did not decrease below this threshold. In this case, we assigned the time it took to reach the minimum body temperature ('Time to Min. T_b') to be 24 h, indicating that the mouse did not decrease its body temperature below this required threshold during the 24-h recording.

Identification of *Adcyap1*$^+$ cells that express TeLC–eYFP

Sample preparation. TeLC-silenced Adcyap1-2A-Cre mice were euthanized and the brains were processed as indicated in section 'Combined fluorescence in situ hybridization and immunofluorescence' using the following reagents: for the detection of TeLC–eYFP$^+$ cells 1:1,000 chicken anti-GFP antibody (Abcam, ab13970) was used as the primary antibody and 1:500 donkey anti-chicken 488 antibody (Jackson ImmunoResearch Laboratories, 703-545-155) as the secondary antibody.

Sample imaging. Sections containing avMLPA were imaged on a Leica SPE confocal microscope using an ACS APO 20×/0.60 IMM CORR objective (Harvard NeuroDiscovery Center). Tiled MPA areas were imaged with a single optical section to avoid counting the same cell across multiple optical sections. Channels were imaged sequentially to avoid any optical crosstalk.

Image analysis. To determine the fraction of ISH *Adcyap1*$^+$ cells that are co-positive for eYFP, *Adcyap1*$^+$ cells in each image were manually marked while staying blinded to the eYFP$^+$ cells. After the identification of *Adcyap1*$^+$ cells, each cell was evaluated for whether it appeared positive for eYFP.

Fibre photometry

Set-up. A three-channel multi-fibre photometry system (Neurophotometrics Ltd) was used for these experiments. In brief, light from three LEDs of different wavelengths (470 nm and 560 nm in phase, and 415 nm out of phase) were bandpass filtered and directed down a fibre-optic patch cord via a 20× objective. This was coupled to a fibre-optic cannula implanted in the mouse. Fluorescence emission from GCaMP6s and mCherry was collected through the same cannula and patch cord, split by a 532-nm longpass dichroic, bandpass filtered, and focused onto opposite sides of a CMOS camera sensor.

Data were acquired and quantified using the open-source software Bonsai[67] by drawing a region of interest around the two images (green and red) of the patch cord and calculating the mean pixel value. To perform longitudinal fibre photometry recordings, the duty cycle of the excitation light was decreased to 10% (interleaved 470 nm + 560 nm / 415 nm with 25 ms period at 4 Hz). LED light was delivered at the minimum power and resulted in about 15 µW of 470 nm light and 25 µW of total light at the tip of the patch cord. A pigtailed fibre-optic rotary joint (Doric FRJ_1x1_PT_200/220/LWMJ-0.37_1.0_FCM_0.15_FCM) was connected to the patch cord (Doric MFP_200/220/900-0.37_#.#_FC_MF1.25) to eliminate bending and coiling of the patch cord.

Fasting and CHA-administration photometry session. Mice were placed in custom-built cages to allow free movement during the entire recording session in the dark. The cages were placed onto radiofrequency receiver platforms (Starr Life Science ER4000). Core body temperature and gross motor activity were logged every 10 s. For fasting-induced torpor sessions, water and nesting material only were provided. The fasting was initiated at the beginning of the dark cycle, while the recordings were started up to several hours after the onset of fasting. For CHA-administration sessions, mice were given access to excess food so as to be maintained in a fed state. CHA (0.2 mg kg^{-1}) was administered via intraperitoneal injection without pausing the recordings.

Analysis of torpor entry, maintenance and arousal for fibre photometry. Within each recording session, a temperature threshold was determined under which the mouse was considered to be in torpor. This threshold was set as 2 °C below the top 95% percentile of the temperatures recorded during the entire session (35.6 ± 0.2 °C, mean ± s.e.m.). Any time the mouse was in torpor (as determined by low core body temperature) and the core body temperature was decreasing by more than 0.05 °C min^{-1} the mouse was considered to be entering torpor (Fig. 4i). Anytime the mouse was in torpor (as determined by low core body temperature) and its core body temperature was increasing by more than 0.1 °C min^{-1} the mouse was considered to be arousing from torpor (Fig. 4j).

Analysis of fibre-photometry data. Custom-written MATLAB software was used to analyse photometry data. Background signal measurement (mean signal with excitation lights off) was first subtracted from all signals. To correct for photobleaching and heat-mediated LED decay, the isosbestic signal was fit with a biexponential that was then linearly scaled to the calcium-dependent fluorescence signal F. To calculate dF/F, we divided the signal by this scaled fit. A three-minute sliding window was applied to calculate the local baseline (10th percentile value) and standard deviation of the dF/F values (Extended Data Fig. 9d). Peaks were chosen on the basis of the prominence (top 1%) of all peaks identified with MATLAB *findpeaks* function. The body temperature data set was linearly interpolated, such that the number of samples was equal to the number of photometry data points.

To plot the distribution of the baseline, peak frequency and standard deviation across non-torpor and different stages of torpor, the entire recording was separated into tiled 3-min periods. For each period the average was calculated, and all the 3-min time periods that overlap with each stage of torpor (torpor entry, torpor, torpor arousal) or non-torpor were plotted as a box plot (mean ± s.e.m., Fig. 4m, n, Extended Data Fig. 9e–g).

Temperature challenge. The temperature challenge was set up and performed similarly to that described previously[16]. Food was provided to the mice in the chamber. The raw calcium-dependent GCaMP6s fluorescence signal was smoothed over a 5-element moving average window, and the baseline F was defined as the average fluorescence of a 10-min window at 25 °C before the first ramp of temperature. dF/F was calculated by dividing the smoothed calcium-dependent GCaMP6s signal by the baseline signal.

Fibre-photometry model. Because we observed marked, statistically significant changes in neural activity when mice were entering and maintaining torpor compared with non-torpid mice or mice that were arousing from torpor, we investigated whether fibre-photometry data are sufficient to determine when mice were entering and maintaining torpor. We first extracted several features from our photometry data—baseline signal, frequency of large peaks and the standard deviation—and calculated the average value for each of these features across a 3, 10 and 30-min sliding window, resulting in a total of 9 distinct features. Using these 9 data features, we performed unsupervised k-means clustering across each of the recording sessions ($n = 8$). Silhouette scores were used to determine the optimum number of clusters ($n = 2$), suggesting that fibre photometry data during each recording session could be robustly grouped into two clusters. Next, we asked which of the features that were included in the k-means clustering contributed most to these clusters and observed that the standard deviation and the baseline calculated from the 3-min sliding window were the main contributors. We therefore used these two features to cluster our fibre-photometry data and investigate whether the clusters (states) accurately correlate with the behavioural data for when mice are entering or maintaining torpor. To evaluate the specificity and sensitivity of our model, we cross-referenced the states (clusters) generated by features of the neural data to torpid versus non-torpid periods as defined by movement and body temperature. The sensitivity of the model that was based on the photometry data was calculated by dividing the amount of time that the model accurately determined torpor entry and maintenance by the total amount of time that a mouse spent entering or maintaining torpor (Extended Data Fig. 9j, k). The model specificity was calculated by dividing the amount of time that the model accurately determined torpor entry or maintenance by the total amount of time that the model calculated that the mouse would be entering or maintaining torpor (whether it was accurate or not, Extended Data Fig. 9j, k). To investigate what the accuracy of the model would have been by chance, we randomly shuffled the model output for each of the recording sessions and evaluated the sensitivity and specificity of this shuffled model (Extended Data Fig. 9k).

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

RNA-sequencing data have been deposited in the Gene Expression Omnibus with accession number GSE149344. Additional data supporting the findings of this study are available from the corresponding authors upon reasonable request.

Code availability

Custom code used in this study is available from the corresponding authors upon reasonable request.

51. Alexander, G. M. et al. Remote control of neuronal activity in transgenic mice expressing evolved G protein-coupled receptors. *Neuron* **63**, 27–39 (2009).
52. Cicconet, M. & Hochbaum, D. R. A. Supervised, symmetry-driven, GUI toolkit for mouse brain stack registration and plane assignment. Preprint at https://www.biorxiv.org/content/10.1101/781880v1 (2019).
53. Lein, E. S. et al. Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168–176 (2007).
54. Oh, S. W. et al. A mesoscale connectome of the mouse brain. *Nature* **508**, 207–214 (2014).
55. Renier, N. et al. Mapping of brain activity by automated volume analysis of immediate early genes. *Cell* **165**, 1789–1802 (2016).
56. Zhang, Y. et al. Identifying local and descending inputs for primary sensory neurons. *J. Clin. Invest*. **125**, 3782–3794 (2015).
57. Klein, A. M. et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**, 1187–1201 (2015).
58. Macosko, E. Z. et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
59. Stroud, H. et al. Early-life gene expression in neurons modulates lasting epigenetic states. *Cell* **171**, 1151–1164 (2017).
60. Hrvatin, S. et al. A scalable platform for the development of cell-type-specific viral drivers. *eLife* **8**, e48089 (2019).
61. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst*. **8**, 281–291 (2019).
62. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol*. **36**, 411–420 (2018).

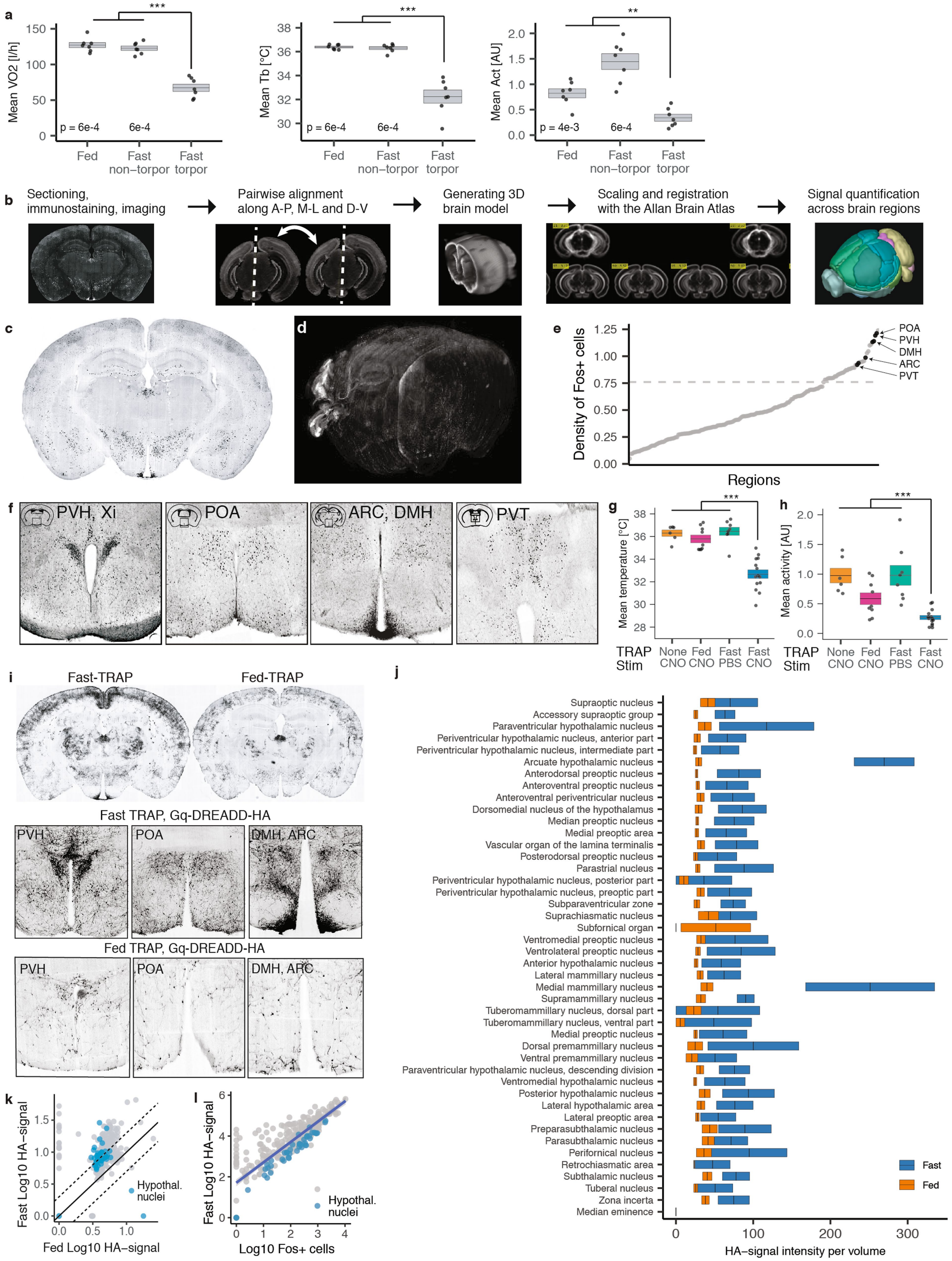
63. Ilicic, T. et al. Classification of low quality cells from single-cell RNA-seq data. *Genome Biol*. **17**, 29 (2016).
64. van den Brink, S. C. et al. Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nat. Methods* **14**, 935–936 (2017).
65. Hrvatin, S. et al. Single-cell analysis of experience-dependent transcriptomic states in the mouse visual cortex. *Nat. Neurosci*. **21**, 120–129 (2018).
66. Buettner, F. et al. Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nat. Biotechnol*. **33**, 155–160 (2015).
67. Lopes, G. et al. Bonsai: an event-based framework for processing and controlling data streams. *Front. Neuroinform*. **9**, 7 (2015).

Acknowledgements We thank B. Sabatini, M. Andermann, B. Lowell, C. Saper and S. R. Datta for feedback on this work; members of the Sabatini and W. Regehr laboratories for reagents; members of the Greenberg, Sabatini, Andermann, Lowell and Datta laboratories for discussions; and A. Mina for technical assistance. This project relied on the Beth Israel Deaconess Medical Center Metabolic Core for experiments in metabolic cages, Boston Children's Hospital Viral Core for AAV packaging, the Harvard NeuroDiscovery Center Enhanced Neuroimaging Core and the Neurobiology Imaging Facility (NINDS P30 Core Center grant NS072030) for imaging, the Data Analysis Core at Harvard Medical School for help with image analysis, and the Research Instrumentation Core Facility for engineering support. S.S. acknowledges support from a Herchel Smith Fellowship. This work was supported by National Institutes of Health grants R01 NS028829 and BRAIN Initiative grant R01 MH114081 to M.E.G., R01 DK107717 to A.S.B. and the Warren Alpert Distinguished Scholar Award to S.H.

Author contributions S.H. conceived the study and designed, performed and analysed experiments. S.S. designed, performed and analysed experiments. O.F.W., H.Y., A.J.L.-P., E.G.A. and M.E.P. performed and analysed experiments. S.A. advised on and analysed fibre photometry experiments. M.C. wrote the code to register and analyse images of brain sections. A.S.B., E.C.G. and M.E.G. advised on the study. S.H., E.C.G. and M.E.G. obtained funding for the research. S.H., S.S., E.C.G. and M.E.G. wrote the manuscript.

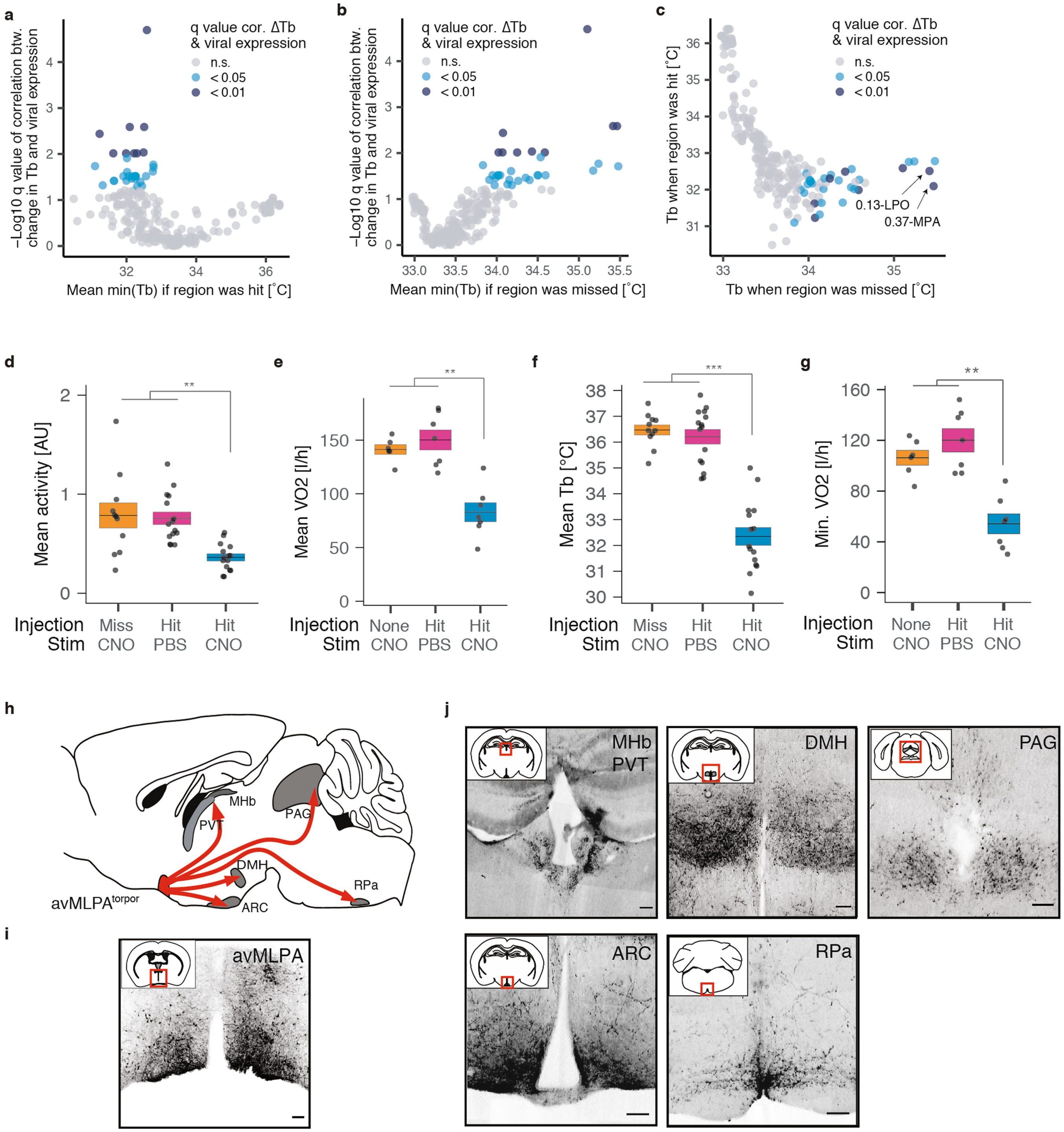
Competing interests S.A. is the founder and CEO of Neurophotometrics Ltd., which manufactures fibre photometry systems. All other authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2387-5.
Correspondence and requests for materials should be addressed to S.H. or M.E.G.
Peer review information *Nature* thanks Trygve E. Bakken, Shaun F. Morrison and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.
Reprints and permissions information is available at http://www.nature.com/reprints.



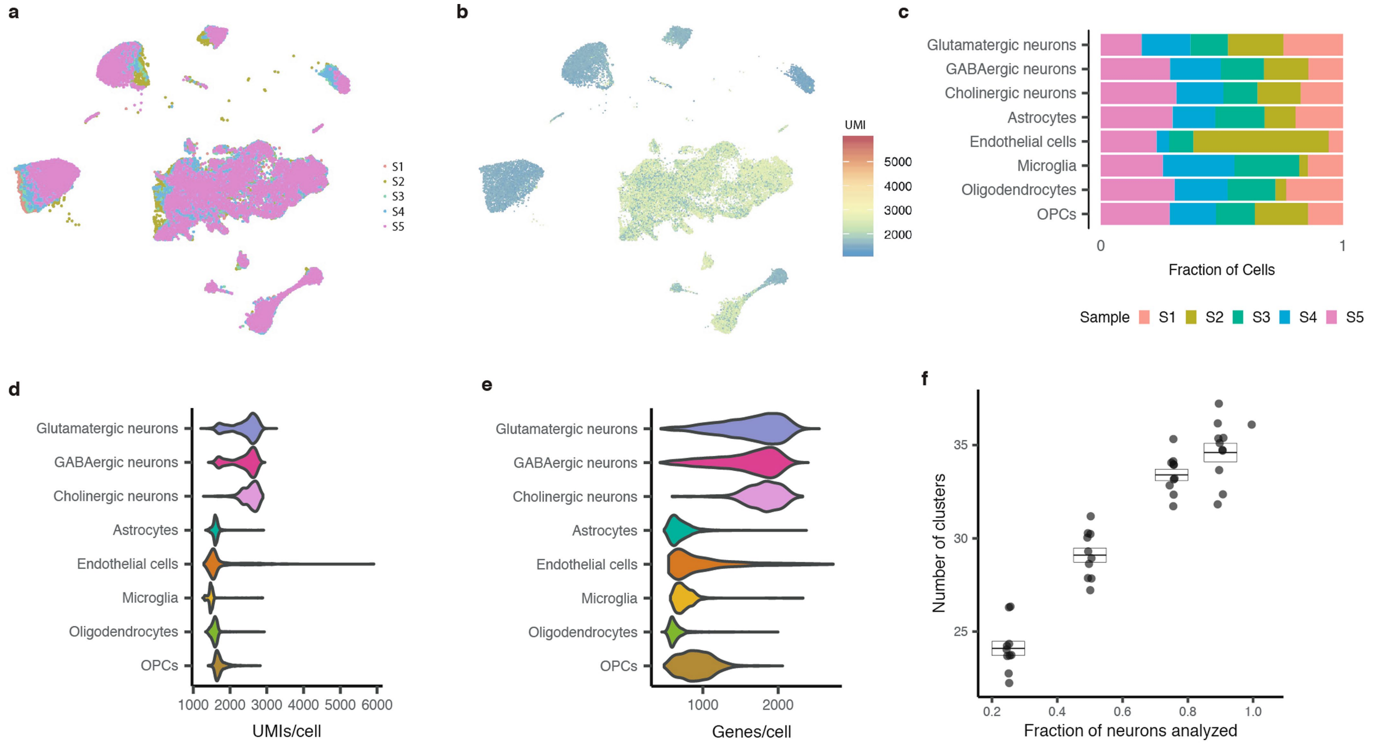
Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Torpor metabolic rate, brain-wide search for torpor-regulating cells and chemogenetic reactivation of FosTRAP-Gq mice. **a**, Mean metabolic rate (VO_2), body temperature (T_b) and gross motor activity (Act) of mice in torpor compared to mice that are fed or fasted yet not in torpor ($n = 7$, *P* values indicated on the graph). **b**, Schematic for the whole-brain reconstruction of FOS staining. **c**, Example brain slice showing Fos staining in a fasted torpid mouse (representative of $n = 3$ mice). **d**, 3D-reconstructed Fos-stained brain slices from a fasted torpid mouse. **e**, Average density of FOS^+ cells (number of cells divided by the volume of the region, $n = 3$ mice, see Methods) across 179 brain regions that had on average at least 100 FOS^+ cells. Paraventricular hypothalamus (PVH), a subregion of the preoptic area (POA), arcuate nucleus (ARC), dorsomedial hypothalamus (DMH) and paraventricular thalamus (PVT) are indicated. **f**, FOS staining of the PVH, xiphoid nucleus (Xi), POA, ARC, DMH and PVT of fasted torpid mice ($n = 3$ mice). **g**, Mean core body temperature (T_b) over 4 h after CNO administration is significantly lower in torpor-TRAP ($n = 14$ mice) compared to non-TRAP ($n = 6$ mice, $P = 5.2 \times 10^{-5}$) and fed-TRAP ($n = 9$ mice, $P = 2.9 \times 10^{-5}$) mice and compared to torpor-TRAP mice injected with PBS ($n = 8$ mice, $P = 2.5 \times 10^{-5}$). **h**, Mean activity over 4 h after CNO administration is significantly lower in torpor-TRAP ($n = 14$ mice) compared to non-TRAP ($n = 6$ mice, $P = 5.2 \times 10^{-5}$) and fed-TRAP ($n = 9$ mice, $P = 9.8 \times 10^{-3}$) mice and compared to torpor-TRAP mice injected with PBS ($n = 8$ mice, $P = 2.5 \times 10^{-5}$). **i**, Coronal brain sections from FosTRAP, LSL-Gq-DREADD-HA mice TRAPed during fasting-induced torpor (fast-TRAP, $n = 2$ mice) or in a fed state (fed-TRAP, $n = 4$ mice) and immunostained for HA. Staining in selected brain areas (PVH, POA, ARC, DMH and PVT) is shown. **j**, Volume-normalized signal intensity of HA staining across different hypothalamic nuclei in four fed-TRAP and two fast-TRAP mice. **k**, Brain-wide quantification of HA staining from four fed-TRAP and two fast-TRAP mice. Numerous (190/316) brain regions, including 32 hypothalamic areas, show increased Gq-DREADD-HA expression (>2-fold) in fast-TRAP mice compared to fed-TRAP mice. The solid line indicates unity, dashed lines indicate twofold differences. **l**, Correlation across brain regions between the number of FOS^+ cells in torpid mice and the levels of Gq-DREADD expression in fast-TRAP mice ($R = 0.83$, $P = 2.2 \times 10^{-16}$, Pearson correlation test, $n = 316$ regions). All box plots indicate mean ± s.e.m. All *P* values are calculated using two-tailed Mann–Whitney *U*-tests, $^{**}P < 0.01$, $^{***}P < 0.001$.

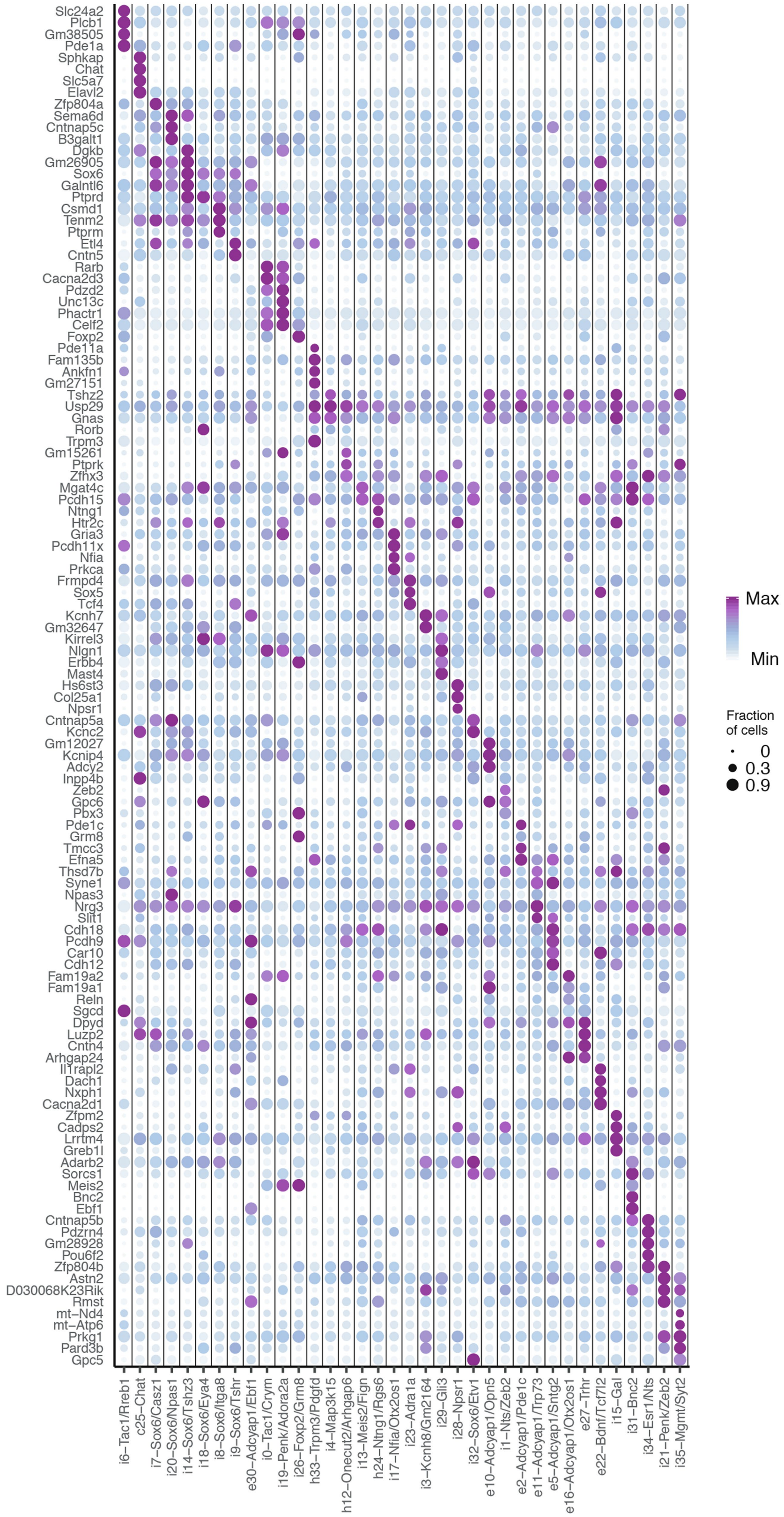


Extended Data Fig. 2 | See next page for caption.

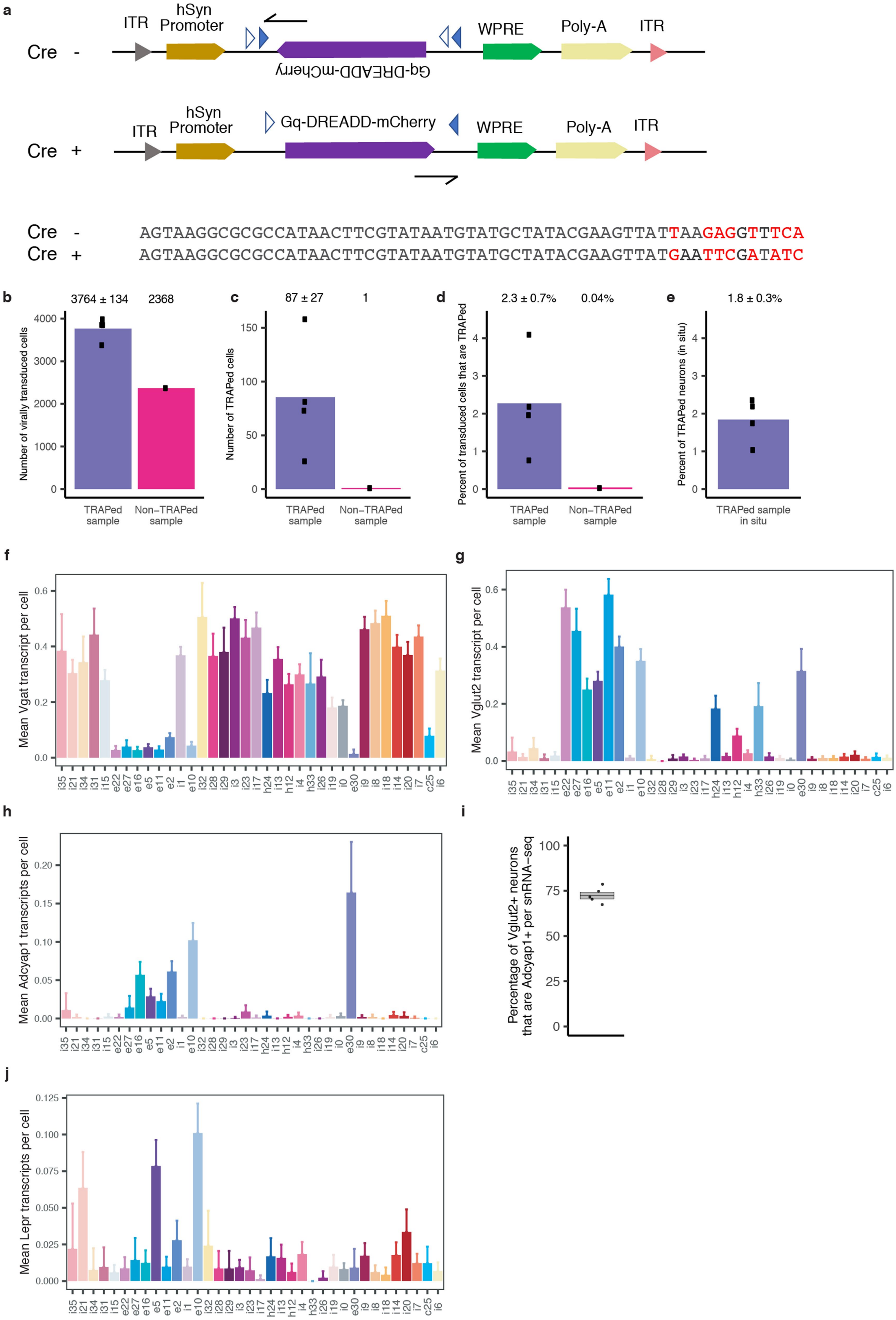
Extended Data Fig. 2 | Chemogenetic reactivation of torpor-TRAPed neurons in different hypothalamic regions and anterograde projections of torpor-TRAPed avMLPA neurons. a–c, AAV-DIO-Gq–mCherry was injected into different hypothalamic regions of FosTRAP mice ($n = 54$ mice). After TRAPing during torpor, we administered CNO and measured the effect of the reactivation of torpor-TRAPed neurons within the virally injected region on core body temperature. All mice were euthanized, and the expression of the virally derived Gq-DREADD–mCherry was evaluated in each mouse across 277 hypothalamic nuclei. **a**, **b**, Each circle represents one of the 277 hypothalamic nuclei, and the *y* axis represents the $-\log_{10}$ FDR-corrected *q* value of the Pearson correlation (across 54 mice, *q* values displayed in Supplementary Table 3) between the viral expression in that nucleus and the decrease in T_b that was observed after CNO stimulation. Next, for each nucleus, 54 mice were grouped into those in which the nucleus was hit (**a**) and those in which it was missed (**b**). For each of the two groups of mice, the minimum body temperature after CNO administration was averaged and plotted. **c**, For each nucleus and the corresponding two groups of mice, the minimum body temperature after CNO administration was plotted (hit group, *y* axis; missed group, *x* axis). Arrows indicate anterior MPA and LPO regions. When these regions were hit with the virus and the TRAPed neurons were chemogenetically reactivated, the body temperature of the mouse decreased, whereas when these regions were missed the body temperature did not decrease. **d**, Mean activity over 4 h after CNO administration is significantly lower in avMLPA-hit ($n = 15$ mice) compared with avMLPA-missed ($n = 11$ mice, $P = 1.4 \times 10^{-3}$) mice, and compared with avMLPA-hit mice injected with PBS ($n = 15$ mice, $P = 4.8 \times 10^{-6}$). **e**, Mean metabolic rate (VO_2) over 4 h after CNO administration is significantly lower in avMLPA-hit ($n = 7$ mice) compared with non-injected ($n = 6$ mice, $P = 2.3 \times 10^{-3}$) mice or avMLPA-hit mice injected with PBS ($n = 6$ mice, $P = 1.2 \times 10^{-3}$). **f**, Mean core body temperature (T_b) over 4 h after CNO administration is significantly lower in avMLPA-hit ($n = 15$ mice) compared to avMLPA-missed ($n = 11$ mice, $P = 2.6 \times 10^{-7}$) mice or avMLPA-hit mice injected with PBS ($n = 15$ mice, $P = 9.0 \times 10^{-8}$). **g**, Minimum metabolic rate (VO_2) over 4 h after CNO administration is significantly lower in avMLPA-hit ($n = 7$ mice) compared with non-injected ($n = 6$ mice, $P = 2.3 \times 10^{-3}$) mice or avMLPA-hit mice injected with PBS ($n = 6$ mice, $P = 1.2 \times 10^{-3}$). **h**, Schematic showing projections of TRAPed avMLPA$^{\text{torpor}}$ neurons. **i**, **j**, Gq-DREADD–mCherry fusion protein expression was used to visualize the projection of TRAPed avMLPA$^{\text{torpor}}$ neurons across the brain ($n = 4$ mice). **i**, Expression of mCherry near the injection site (avMLPA). **j**, Representative images of projections to the medial habenula (MHb), PVT, DMH, periaqueductal grey (PAG), ARC and raphe pallidus (RPa). Scale bars, 50 µm. For the box plots, the centre line and box boundaries indicate mean ± s.e.m. All *P* values were calculated using a two-tailed Mann–Whitney *U*-test, $^{**}P < 0.01$, $^{***}P < 0.001$.



Extended Data Fig. 3 | snRNA-seq metrics. a, **b**, UMAP plot of 39,562 nuclei from the avMLPA of five mice, in which the colours denote cells derived from each mouse (**a**) or the number of unique transcripts (UMI) per nucleus (**b**). **c**, Relative contribution of each sample ($n = 5$ mice) towards the total cell population making up each main cell class. **d**, Violin plot of the distribution of UMIs per cell for each main cell class (glutamatergic neurons, $n = 11{,}275$ cells; GABAergic neurons, $n = 16{,}307$ cells; cholinergic neurons, $n = 521$ cells; astrocytes, $n = 3{,}479$ cells; endothelial cells, $n = 421$ cells; microglia, $n = 1{,}247$ cells; oligodendrocytes, $n = 4{,}718$ cells; and OPCs, $n = 1{,}594$ cells). **e**, Violin plot of the distribution of genes per cell for each main cell class. **f**, Number of neuronal clusters formed when different fractions (25%, 50%, 75% and 90%) of total neurons ($n = 7{,}025$, $14{,}051$, $21{,}077$ and $25{,}292$, respectively) are used for clustering. For each fraction a random subset of neurons was used and the analysis was repeated ten times. For the box plots, the centre line and box boundaries indicate mean ± s.e.m, and the violin plot shows the distribution from the lowest to the largest value.

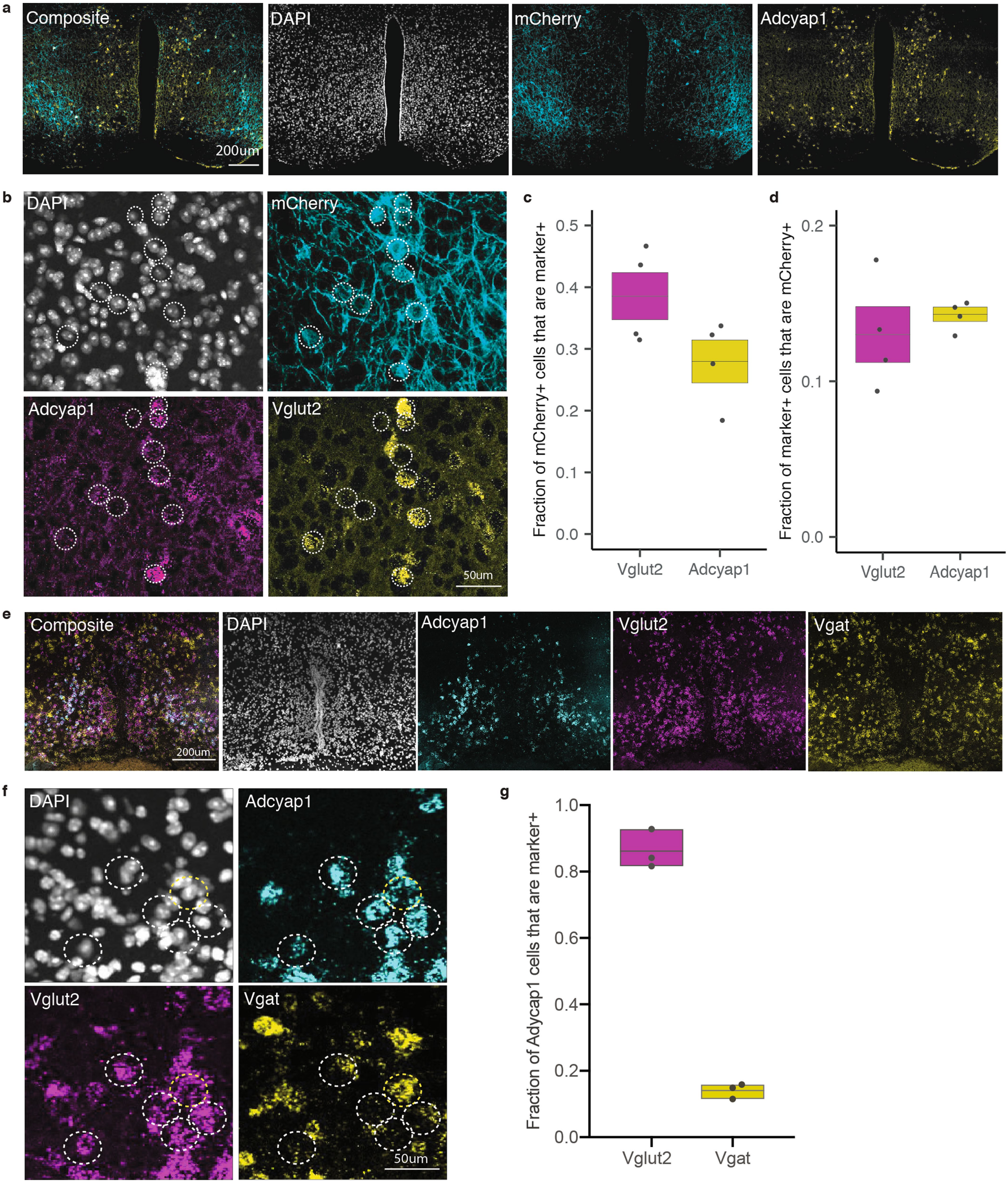


Extended Data Fig. 4 | Marker gene expression across neuronal cell types. The colour denotes mean expression across all nuclei normalized to the highest mean across cell types, and the size represents the fraction of nuclei in which the marker gene was detected. Cell types are organized on the basis of hierarchical clustering across all variable genes. The five most unique makers are identified and plotted for each cell type unless a marker was identified across multiple cell types, in which case it was plotted only once.



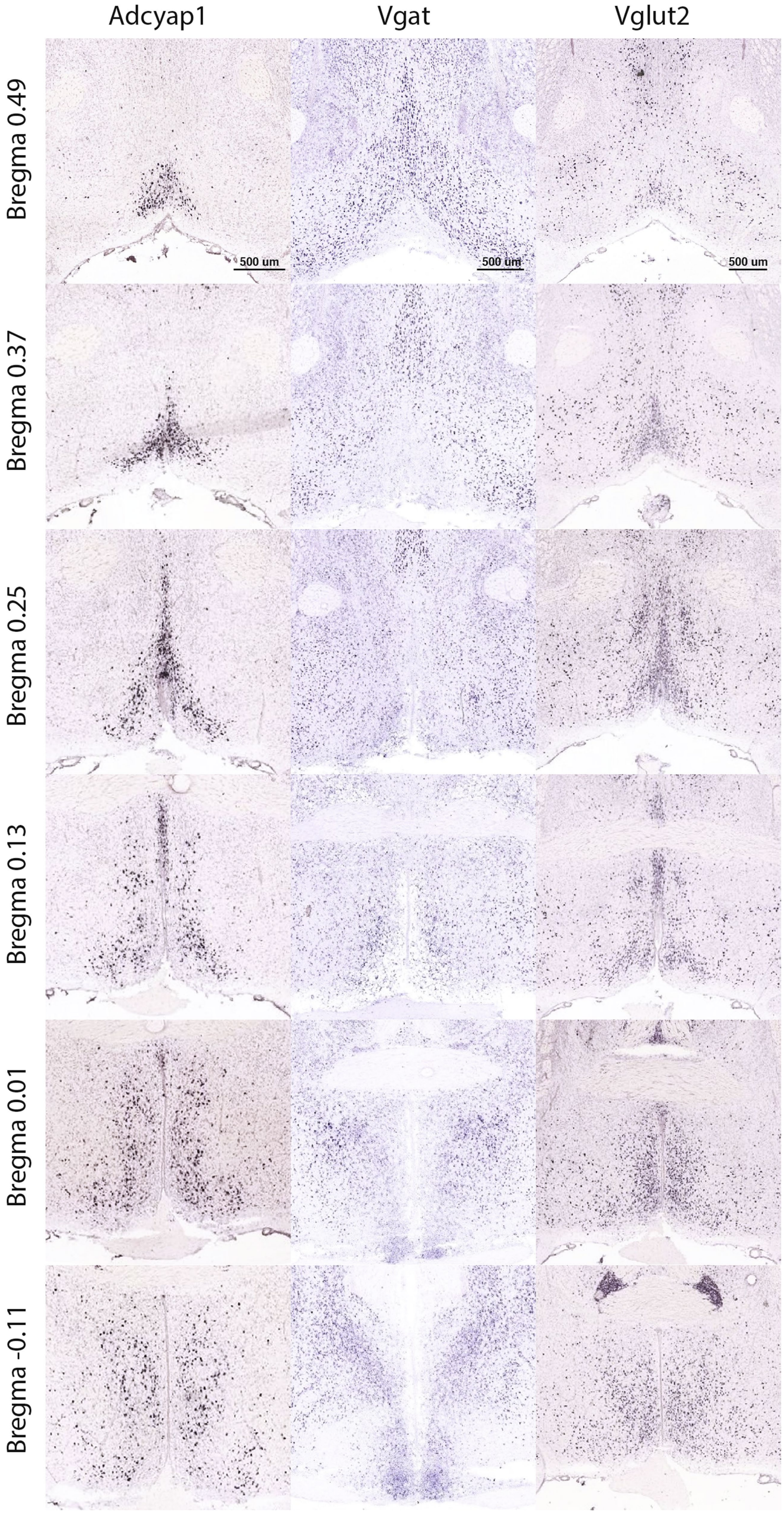
Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Strategy for identifying TRAPed torpor-regulating neurons via snRNA-seq and gene expression of marker genes in the avMLPA. a, Schematic for the identification of Cre-dependent AAV-DIO-Gq-DREADD–mCherry mRNA with or without recombination. Top, AAV-DIO-Gq-DREADD–mCherry vector map before (Cre$^-$) and after (Cre$^+$) Cre-mediated recombination. Blue and white triangles surrounding the Gq-GREADD–mCherry indicate *loxP* sites. Black arrows indicate the binding site of the sequencing primer. ITR, inverted terminal repeats; WPRE, Woodchuck Hepatitis Virus post-transcriptional regulatory element; Poly-A, polyadenylation signal. Bottom, owing to the Cre-mediated inversion in the AAV-DIO-Gq-DREADD–mCherry vector, the mRNA transcript sequence 3′ of the sequencing primer is different after Cre-mediated recombination, enabling us to identify TRAPed cells during snRNA-seq as those cells in which the viral mRNA contains the recombined (Cre$^+$) sequence. **b**, Quantification of the number of virally transduced cells in TRAPed (*n* = 4 mice) and non-TRAPed (*n* = 1 mouse) samples. **c**, Quantification of the number of TRAPed cells in TRAPed (86 ± 27 cells) and non-TRAPed (1 cell) samples. **d**, The percentage of transduced cells that are TRAPed in TRAPed (2.3 ± 0.7%, *n* = 4 mice) and non-TRAPed (0.04%, *n* = 1 mouse) samples based on snRNA-seq analysis. **e**, The percentage of TRAPed neurons in TRAPed samples (1.8 ± 0.3%, *n* = 4 mice) based on fluorescence in situ hybridization analysis. **f**–**h**, Mean transcripts per cell across all neuronal cell types identified in snRNA-seq for *Vgat* (*Slc32a1*, marker of GABAergic neurons) (**f**), *Vglut2* (*Slc17a6*, marker of glutamatergic neurons) (**g**) and *Adcyap1* (adenylate cyclase-activating peptide 1) (**h**). **i**, snRNA-seq indicates that e2, e5, e10, e11, 16 and e30 represent *Vglut2*$^+$*Adcyap1*$^+$ cell types, whereas e22, e27, h33, h12 and h24 are *Vglut2*$^+$*Adcyap1*$^-$. On the basis of this categorization, 72.4 ± 2.2% of *Vglut2*$^+$ neurons are *Adcyap1*$^+$ (*n* = 5 mice). For the box plot, the centre line and box boundaries indicate mean ± s.e.m. **j**, Mean transcripts per cell across all neuronal cell types identified in snRNA-seq for *Lepr*. *Adcyap1*$^+$ clusters e5 and e10 express *Lepr*. Data are mean ± 2s.e.m.

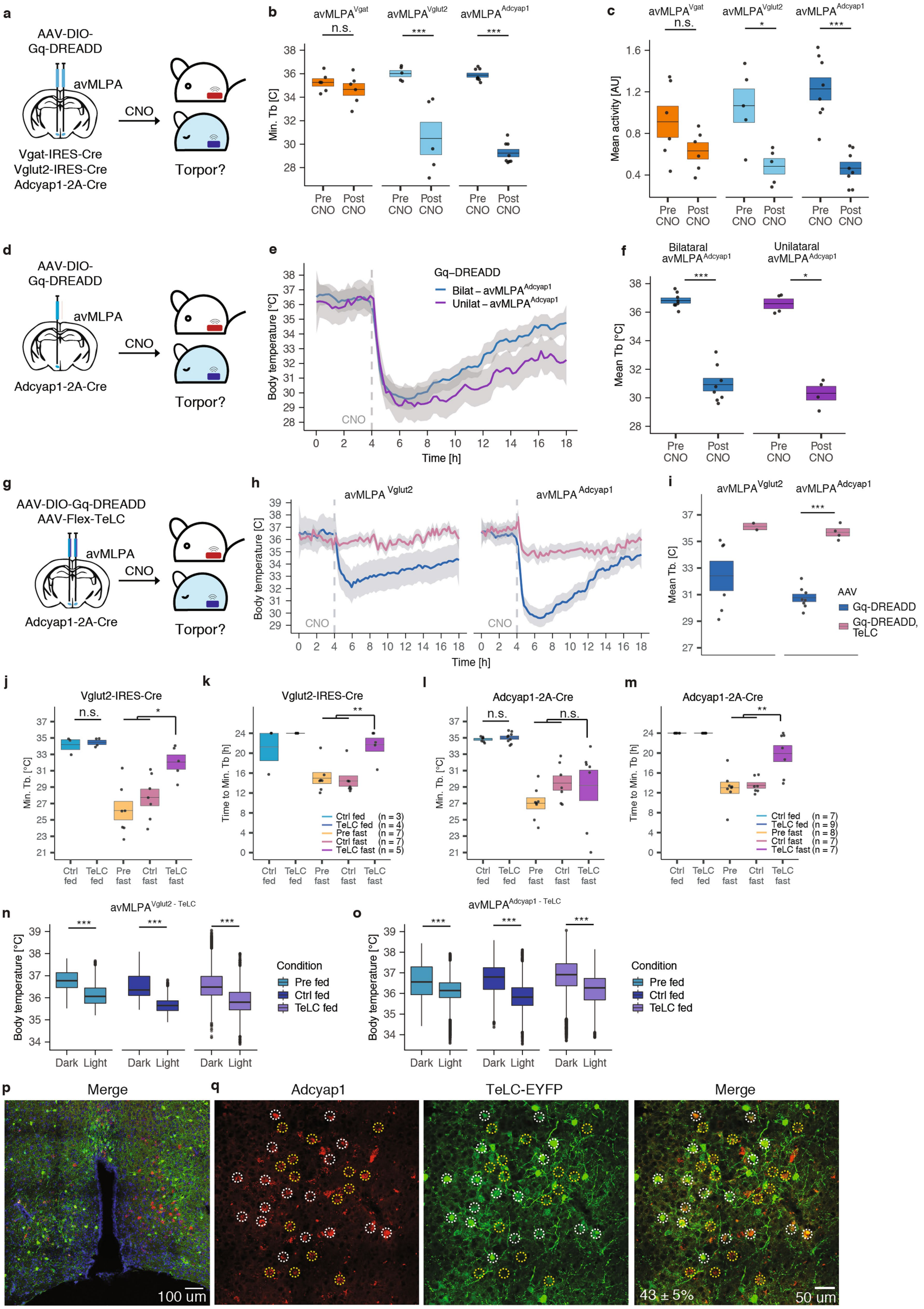


Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | In situ hybridization analysis of torpor-regulating avMLPA neurons. **a**, Coronal sections showing the avMLPA of FosTRAP mice ($n = 4$ mice) injected with AAV-DIO-Gq-DREADD–mCherry and torpor-TRAPed. Immunofluorescent staining against mCherry indicates the location of avMLPA$^{\mathrm{torpor}}$ neurons (cyan), whereas in situ hybridization indicates the expression of the marker gene *Adcyap1*. **b**, High-magnification images of staining shown in **a** indicate the location of mCherry^{+} avMLPA$^{\mathrm{torpor}}$ neurons (cyan), whereas in situ hybridization indicates the expression of marker genes *Adcyap1* and *Vglut2*. Example avMLPA$^{\mathrm{torpor}}$ mCherry^{+} cells are circled. Several mCherry^{+} cells express *Adcyap1* and/or *Vglut2*. **c**, Quantification of the fraction of avMLPA$^{\mathrm{torpor}}$ neurons that express *Adcyap1* (28.8 ± 3.5%, $n = 4$ mice) and *Vglut2* (38.5 ± 3.8%, $n = 4$ mice). **d**, Quantification of the fraction of avMLPA$^{\mathrm{Adcyap1+}}$ (14.3 ± 0.5%, $n = 4$ mice) and avMLPA$^{\mathrm{Vglut2+}}$ (13.0 ± 1.8%, $n = 4$ mice) neurons that are torpor-TRAPed. **e**, Coronal section showing the avMLPA of FosTRAP mice. In situ hybridization shows cells that are positive for *Adcyap1* (cyan), *Vgat* (yellow) and *Vglut2* (purple). The composite image indicates co-expression of multiple markers. **f**, High-magnification image with example *Adcyap1*$^{+}$ cells circled. White circles indicate *Adcyap1*$^{+}$ cells that are positive for *Vglut2* and negative for *Vgat*, whereas yellow circles indicate all *Adcyap1*$^{+}$ cells that are positive for *Vgat* (even if co-positive with *Vglut2*). **g**, The fraction of *Adcyap1*$^{+}$ cells that are positive for *Vglut2* or *Vgat* (82 ± 3% or 14 ± 1%, respectively, $n = 3$ mice). For the box plots, the centre line and box boundaries indicate mean ± s.e.m.

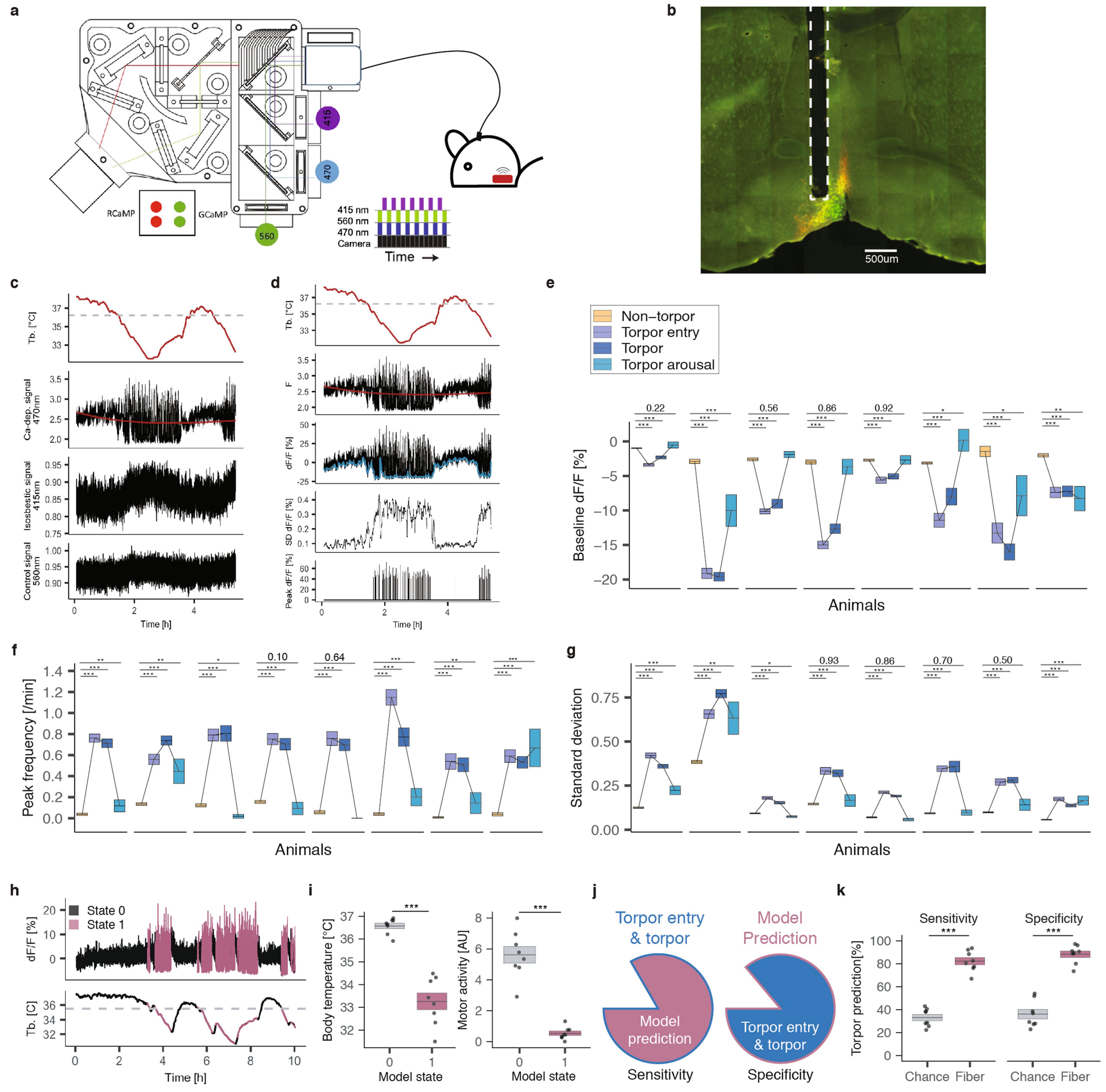


Extended Data Fig. 7 | Expression pattern of *Vgat*, *Vglut2* and *Adcyap1* in the anterior POA. Coronal sections adapted from the Allen Mouse Brain Atlas[32]. Anterior–posterior coordinates relative to bregma are indicated for each set of images.



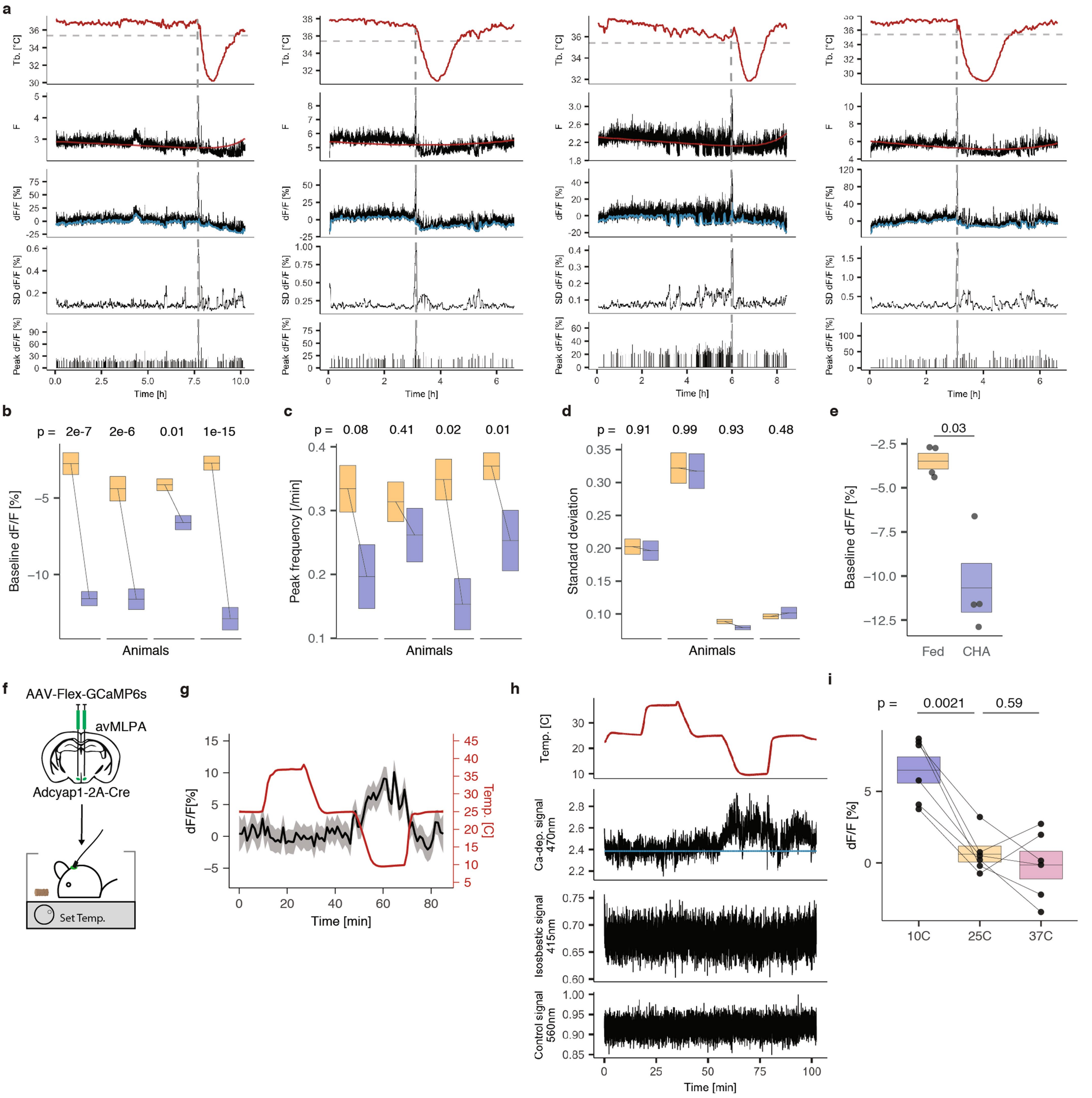
Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Chemogenetic stimulation and silencing of avMLPA$^{\mathrm{Vgat}}$, avMLPA$^{\mathrm{Vglut2}}$ or avMLPA$^{\mathrm{Adcyap1}}$ neurons. a–**c**, Stereotaxic viral injection of AAV-DIO-Gq-DREADD and subsequent chemogenetic stimulation of avMLPA$^{\mathrm{Vgat}}$ ($n = 6$ mice), avMLPA$^{\mathrm{Vglut2}}$ ($n = 5$ mice) or avMLPA$^{\mathrm{Adcyap1}}$ ($n = 8$ mice) neurons. **a**, Experimental schematic. **b**, Minimum core body temperature of avMLPA$^{\mathrm{Vgat}}$ mice (orange, $P = 0.48$), avMLPA$^{\mathrm{Vglut2}}$ mice (light blue, $P = 8 \times 10^{-3}$) and avMLPA$^{\mathrm{Adcyap1}}$ mice (dark blue, $P = 1.6 \times 10^{-4}$) before and after chemogenetic stimulation with CNO. **c**, Mean activity of the same avMLPA$^{\mathrm{Vgat}}$ mice ($P = 0.24$), avMLPA$^{\mathrm{Vglut2}}$ mice ($P = 0.032$) and avMLPA$^{\mathrm{Adcyap1}}$ mice ($P = 1.6 \times 10^{-4}$), before and after chemogenetic stimulation with CNO. **d**, Schematic showing the unilateral stereotaxic viral injection of AAV-DIO-Gq-DREADD and subsequent chemogenetic stimulation of avMLPA$^{\mathrm{Adcyap1}}$ neurons. **e**, Change in mean core body temperature after bilateral ($n = 8$ mice) and unilateral ($n = 4$ mice) chemogenetic stimulation of avMLPA$^{\mathrm{Adcyap1}}$ neurons. The dashed line indicates CNO administration. Coloured lines indicate the mean core body temperature across mice; grey shading indicates the 95% confidence interval. **f**, Mean core body temperature of mice before and after bilateral ($n = 8$ mice, $P = 1.6 \times 10^{-4}$) or unilateral ($n = 4$ mice, $P = 0.03$) chemogenetic stimulation of avMLPA$^{\mathrm{Adcyap1}}$ neurons. **g**, Schematic for the stereotaxic viral co-injection of AAV-Flex-TeLC and AAV-DIO-Gq-DREADD and subsequent chemogenetic stimulation of avMLPA$^{\mathrm{Vglut2}}$ and avMLPA$^{\mathrm{Adcyap1}}$ neurons. **h**, Changes in mean core body temperature after chemogenetic stimulation of avMLPA$^{\mathrm{Vglut2}}$ and avMLPA$^{\mathrm{Adcyap1}}$ neurons that either express the excitatory Gq-DREADD receptor ($n = 6$ and $n = 8$ mice, respectively) or co-express the Gq-DREADD receptor and TeLC, which inhibits synaptic transmission ($n = 2$ and $n = 4$ mice, respectively). The dashed line indicates CNO administration. Coloured lines indicate the mean core body temperature across mice; grey shading indicates the 95% confidence interval. **i**, Quantification of mean core body temperature over 4 h after chemogenetic stimulation in avMLPA$^{\mathrm{Vglut2}}$ and avMLPA$^{\mathrm{Adcyap1}}$ ($P = 1 \times 10^{-6}$) neurons that either solely express the excitatory Gq-DREADD receptor ($n = 6$ and $n = 8$ mice, respectively) or co-express the Gq-DREADD and TeLC ($n = 2$ and $n = 4$ mice, respectively). **j**–**o**, Stereotactic injection of AAV-Flex-TeLC to inhibit synaptic transmission in avMLPA$^{\mathrm{Vglut2}}$ and avMLPA$^{\mathrm{Adcyap1}}$ neurons. **j**, **k**, Core body temperature of fed and fasted Vglut2-IRES-Cre mice from Fig. 4e (the number of mice in each group is indicated on the graph). **j**, The minimum T_{b} is not significantly different between control-fed and TeLC-fed ($P = 0.72$) mice, but is significantly lower in control-fast ($P = 0.018$), and pre-fast ($P = 0.01$) compared to TeLC-fast mice, suggesting that avMLPA$^{\mathrm{Vglut2}}$ activity is necessary for torpor. **k**, Time needed to reach the minimum body temperature (Fig. 4e) is significantly longer in TeLC-fast compared with either pre-fast or control-fast mice ($P = 9.2 \times 10^{-3}$ for both sets). **l**, **m**, Body temperature of fed and fasted Adcyap1-2A-Cre mice from Fig. 4f (the number of mice in each group is indicated on the graph). **l**, The minimum T_{b} is not significantly different between control-fed and TeLC-fed ($P = 0.41$) mice, or in TeLC-fast compared to control-fast ($P = 0.71$) and pre-fast ($P = 0.19$) mice. **m**, Time needed to reach the minimum body temperature (Fig. 4f) is significantly longer in TeLC-fast compared to pre-fast ($P = 2 \times 10^{-3}$) and control-fast ($P = 7 \times 10^{-3}$) mice. **n**, **o**, Core body temperature (measured in 1-min intervals) of fed mice during the 12-h light and 12-h dark cycle in which avMLPA$^{\mathrm{Vglut2}}$ (**n**) or avMLPA$^{\mathrm{Adcyap1}}$ (**o**) neurons were injected with either AAV-Flex-TeLC (TeLC), a control AAV (control), or remained un-injected (Pre). The core body temperature is significantly different between the dark and light cycle across pre-fed ($n = 3$ mice, $n = 3{,}960$ temperature data points, $P = 2 \times 10^{-16}$), control-fed ($n = 2$ mice, $n = 2{,}640$ temperature data points, $P = 2 \times 10^{-16}$) and TeLC-fed ($n = 5$ mice, $n = 6{,}600$ temperature data points, $P = 2 \times 10^{-16}$) Vglut2-IRES-Cre mice (**n**) as well as pre-fed ($n = 4$ mice, $n = 5{,}280$ temperature data points, $P = 2 \times 10^{-16}$), control-fed ($n = 7$ mice, $n = 9{,}240$ temperature data points, $P = 2 \times 10^{-16}$) and TeLC-fed ($n = 8$ mice, $n = 10{,}560$ temperature data points, $P = 2 \times 10^{-16}$) Adcyap-2A-Cre mice (**o**). In the box plots the centre line denotes the median, the box boundaries mark the interquartile range (IQR) and the whiskers extend to 1.5 × IQR and any data points outside this range. **p**, **q**, Coronal section showing the avMLPA of Adcyap1-2A-Cre mice ($n = 2$ mice) injected with AAV-Flex-TeLC-eYFP. Immunofluorescent staining against eYFP indicates the location of silenced TeLC^{+} neurons (green), whereas in situ hybridization indicates the expression of the *Adcyap1* mRNA. **q**, High-magnification image with example *Adcyap1*$^{+}$ cells circled. White circles indicate *Adcyap1*$^{+}$ cells that co-express TeLC–eYFP ($43 \pm 5\%$, $n = 2$ mice), yellow circles indicate *Adcyap1*$^{+}$ that do not co-express TeLC–eYFP. All P values are calculated using a two-tailed Mann–Whitney U-test. NS indicates not statistically significant, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. In the box plots in **b**–**m**, the centre line and box boundaries indicate mean ± s.e.m.



Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Fibre-photometry set-up, recordings and torpor model. **a**, Schematic showing the fibre-photometry set-up. Three LED lights (415 nm, 470 nm and 560 nm) were used as excitation light sources. For all recordings, 470 nm and 560 nm light sources were driven in phase, with 415 nm driven out of phase (Methods). The emitted signals were detected by a digital camera at the end of a patch cord. **b**, Example coronal brain slice from an Adcyap1-2A-Cre mouse co-injected with AAV-DIO-Gq-DREADD–mCherry and AAV-Flex-GCaMP6s and used for fibre photometry studies ($n = 8$ mice). The white dashed lines indicate the location of the optical fibre. Cells co-expressing GCaMP6s (green) and mCherry (red) appear yellow. **c**, Example fibre-photometry recording (from mouse shown in **b**) showing the core body temperature (top) followed by three different signals (470 nm, 415 nm and 560 nm). Here, the 470-nm signal represents the calcium-dependent GCaMP6s signal, the 415-nm signal represents the Ca^{2+}-independent isosbestic GCaMP6s signal, and the 560-nm signal represents the mCherry signal. The red line indicates the scaled fit of the Ca^{2+}-independent 415-nm signal used to normalize the Ca^{2+}-dependent 470-nm signal for Ca^{2+}-independent changes in signal intensity. Both the 415-nm and 560-nm channels serve as controls for heat-mediated LED decay, bleaching of GCaMP6s and movement artefacts. **d**, Recordings of a representative fasting session. Top panel, core body temperature of mice during each recording session (dashed line indicates the threshold body temperature below which the mouse is considered torpid); second panel, raw Ca^{2+}-dependent 470-nm GCaMP6s signal (the red line indicates the scaled fit of the Ca^{2+}-independent 415-nm signal used to normalize for bleaching or other Ca^{2+}-independent changes in signal intensity); third panel, dF/F value relative to the Ca^{2+}-independent scaled fit (blue line indicates the local baseline, which is determined as the 10th percentile of the dF/F value within a sliding three-minute interval); fourth panel, the standard deviation of the dF/F value calculated within a sliding three-minute interval; bottom panel, the dF/F values of the most prominent peaks identified (top 1% of all peaks in the session). **e**–**g**, Quantification of baseline dF/F (%) (**e**), peak frequency (per min) (**f**) and standard deviation (**g**) for non-torpid (yellow), torpor entry (light blue), torpor (blue) and torpor arousal (teal) in 8 individual mice across all 3-min intervals (left to right: $n = 251$, 62, 97, 17, 321, 46, 57, 15, 203, 52, 39, 19, 269, 44, 66, 18, 141, 30, 59, 2, 250, 57, 42, 5, 43, 31, 23, 7, 80, 44, 51, 8 time intervals). In box plots, the centre line and box boundaries indicate mean ± s.e.m. P values greater than 0.05 are indicated. **h**, Example fibre-photometry signal (top) clustered into two states and coloured by state. State 0 corresponds to the mouse being out of torpor or exiting torpor, whereas state 1 corresponds to the mouse entering or maintaining torpor. **i**, Core body temperature (left) and motor activity (right) are significantly lower during state 1 compared with state 0 of the photometry-based model ($n = 8$ mice, $P = 1.6 \times 10^{-4}$). **j**, The time that a mouse spent in torpor (entry or maintenance) was accurately calculated by the model based on the photometry data 82.3 ± 3.2% of the time (model sensitivity). Conversely, whenever the model determined that the mouse was entering or maintaining torpor, its estimation was 88.4 ± 2.8% accurate (specificity). **k**, Model sensitivity and specificity were significantly lower ($P = 1.6 \times 10^{-4}$, $n = 8$ mice) when the temporal relationship between the temperature and the fibre-photometry data was removed. In box plots, the centre line and box boundaries indicate mean ± s.e.m. All P values were calculated using a two-tailed Mann–Whitney U-test. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.



Extended Data Fig. 10 | Fibre-photometry recordings of avMLPA$^{\mathrm{Adcyap1}}$ neurons in fed freely moving mice with CHA-induced hypothermia and changes in ambient temperature. **a**, Fibre-photometry recording data displayed as in Extended Data Fig. 9d. The dashed line indicates the time of CHA administration. **b**–**d**, Baseline d*F*/*F* (**b**), peak frequency (**c**) and standard deviation (**d**) measured for each mouse before and after CHA administration across all recorded three-minute intervals (left to right: $n = 69, 17, 63, 23, 140, 24, 161$ and 37 time intervals). **e**, The mean baseline decreases after CHA treatment ($P = 0.03$, $n = 4$ mice). **f**, Schematic showing the fibre-photometry recording of avMLPA$^{\mathrm{Adcyap1}}$ neurons when mice are exposed to different environmental temperatures with food provided in the chamber. **g**, Mean GCaMP6s signal ($n = 6$ mice) of avMLPA$^{\mathrm{Adcyap1}}$ neurons with environmental temperature changes along a programmed sequence: 25 °C → 37 °C → 25 °C → 10 °C → 25 °C. Grey shading indicates the 95% confidence interval. **h**, Example fibre-photometry recording showing the ambient (chamber) temperature (top) followed by three different signals (470 nm, 415 nm and 560 nm). Signals from the 415-nm and 560-nm channels are used as controls for any potential effects of temperature on the photometry signal. **i**, Mean neuronal responses at different ambient temperatures. avMLPA$^{\mathrm{Adcyap1}}$ neurons are not sensitive to increases in the ambient temperature to 37 °C ($P = 0.59$), and instead appear to be sensitive to a decrease in environmental temperature ($n = 6$ mice, $P = 0.0021$). In box plots, the centre line and box boundaries indicate mean ± s.e.m. All *P* values were calculated using a two-tailed Mann–Whitney *U*-test. NS indicates not statistically significant, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.
]

nature research

Corresponding author(s): Michael Greenberg, Sinisa Hrvatin

Last updated by author(s): May 1, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☐	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	VitalView 5.1 (Starr Life Sciences) for collection of body temperature and activity. Oxymax (Columbus Instruments) for CLAMS experiments. Bonsai 2.4 (open-source) for collection of fiber photometry. Olyvia 2.4 (Olympus) for imaging brain sections. LasX 3.3.0 (Leica) for imaging in situs.
Data analysis	Most of the analysis was performed in R. Image processing was performed in Fiji 2.0.0 (converstion to TIFF), Matlab R2019b (MathWorks) and using Riffle Shuffle (github.com/hms-idac/RiffleShuffle). Single nucleus RNA-seq analysis was performed using Cellranger 3.1.0 (10X Genomics), Scublet (python package) and Seurat 3.1 (R package).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the findings of this study are available from the corresponding author. Raw and processed single-cell RNA-seq counts data and metadata is available at GEO accession GSE149344.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample size. Sample sizes are indicated for each experiment and were chosen based on similar studies.
Data exclusions	Pre-established criteria were used for data inclusion/exclusion. Due to variability across stereotactic injections, only FosTRAP animals that showed a decrease in core body temperature following chemogenetic stimulation were used for snRNA-seq and in situ hybridization. Cell doublets were removed using criteria that is consistent with other publications and is reported in the methods. For fiber photometry experiments, only animals for which we confirmed the correct placement of the fiber were included.
Replication	Each experiment was performed across several animals (numbers indicated in manuscript). Where possible, data from each individual animal is shown in the manuscript indicating the distribution of the results. Stereotactic surgeries were performed by two separate individuals.
Randomization	Assignment of individual mice to different surgical groups and experimental groups was random.
Blinding	The analysis of viral expression across 54 animals and 277 hypothalamic regions was performed blinded to the effect on core body temperature that was previously observed in each animal. Identification of TRAPed neurons (mCherry+ cells) or marker-expressing cells was each performed while staying blinded to the other analysis to avoid bias.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study	
☐	☒	Antibodies
☒	☐	Eukaryotic cell lines
☒	☐	Palaeontology
☐	☒	Animals and other organisms
☒	☐	Human research participants
☒	☐	Clinical data

Methods

n/a	Involved in the study	
☒	☐	ChIP-seq
☒	☐	Flow cytometry
☒	☐	MRI-based neuroimaging

Antibodies

Antibodies used	1:300 rabbit anti-mCherry antibody (Abcam # ab167453), 1:500 donkey anti-rabbit 568 antibody (Life technologies Cat# AB_2534017), 1:2000 rabbit anti-Fos (Cedarlane # 226003(SY)), 1:1000 rabbit anti-HA (Cell Signaling Technology # 3724S), 1:500 donkey anti-rabbit 647 (Life Technologies # A31573), 1:1000 chicken anti-GFP antibody (Abcam, ab13970), 1:500 donkey anti-chicken 488 antibody (Jackson ImmunoResearch Laboratories, 703-545-155)
Validation	All primary antibodies have been previously used in several publications: ab167453 - PMIDs: 29556030, 30528281 226003(SY) - PMIDs: 31097621, 31376224 3724S - PMIDs: 26743492, 25077630 ab13970 - PMIDs: 30385274, 30559277

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	For initial torpor experiments we used adult (6-10-week-old) C57BL/6J (The Jackson Laboratory, Stock # 000664) mice. To generate FosTRAP-Gq mice we crossed Fos2A-iCreER (TRAP2) (The Jackson Laboratory Stock # 030323) with R26-LSL-Gq-DREADD (The Jackson Laboratory Stock # 026220) and used adult (6-18-week-old) male and female F1 progeny. For viral injections we

	used Fos2A-iCreER (TRAP2) (The Jackson Laboratory Stock # 030323), Adcyap1-2A-Cre (The Jackson Laboratory Stock # 030155), Vglut2-ires-cre (The Jackson Laboratory Stock # 028863) and Vgat-IRES-Cre (The Jackson Laboratory Stock # 028862) mice.
Wild animals	The study did not involve wild animals
Field-collected samples	The study did not involve samples collected in the field
Ethics oversight	Animal experiments were approved by the National Institute of Health and Harvard Medical School Institutional Animal Care and Use Committee, following ethical guidelines described in the US National Institutes of Health Guide for the Care and Use of Laboratory Animals.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Article

Hepatic NADH reductive stress underlies common variation in metabolic traits

https://doi.org/10.1038/s41586-020-2337-2

Received: 22 February 2019

Accepted: 11 March 2020

Russell P. Goodman[1,2], Andrew L. Markhard[1], Hardik Shah[1], Rohit Sharma[1], Owen S. Skinner[1], Clary B. Clish[3], Amy Deik[3], Anupam Patgiri[1], Yu-Han H. Hsu[3,4,5], Ricard Masia[6,7], Hye Lim Noh[8], Sujin Suk[8], Olga Goldberger[1], Joel N. Hirschhorn[3,4,5], Gary Yellen[7], Jason K. Kim[8,9] & Vamsi K. Mootha[1,3,10]✉

The cellular NADH/NAD$^+$ ratio is fundamental to biochemistry, but the extent to which it reflects versus drives metabolic physiology in vivo is poorly understood. Here we report the in vivo application of *Lactobacillus brevis* (*Lb*)NOX[1], a bacterial water-forming NADH oxidase, to assess the metabolic consequences of directly lowering the hepatic cytosolic NADH/NAD$^+$ ratio in mice. By combining this genetic tool with metabolomics, we identify circulating α-hydroxybutyrate levels as a robust marker of an elevated hepatic cytosolic NADH/NAD$^+$ ratio, also known as reductive stress. In humans, elevations in circulating α-hydroxybutyrate levels have previously been associated with impaired glucose tolerance[2], insulin resistance[3] and mitochondrial disease[4], and are associated with a common genetic variant in *GCKR*[5], which has previously been associated with many seemingly disparate metabolic traits. Using *Lb*NOX, we demonstrate that NADH reductive stress mediates the effects of *GCKR* variation on many metabolic traits, including circulating triglyceride levels, glucose tolerance and FGF21 levels. Our work identifies an elevated hepatic NADH/NAD$^+$ ratio as a latent metabolic parameter that is shaped by human genetic variation and contributes causally to key metabolic traits and diseases. Moreover, it underscores the utility of genetic tools such as *Lb*NOX to empower studies of 'causal metabolism'.

NADH and NAD$^+$ are essential redox cofactors that lie at the heart of metabolism. They have a particularly important role in hepatic metabolism[6], although there is conflicting evidence as to their causal contributions and directionality in disease. For example, mice fed a high-fat diet (HFD) develop metabolic abnormalities that can be partially reversed by treatment with nicotinamide riboside (NR), which raises total NAD$^+$ levels[7,8]. By contrast, recent studies have implicated an increase in the hepatic cytosolic NADH/NAD$^+$ ratio in the glucose-lowering effects of metformin[9].

This lack of clarity stems in part from a dearth of tools with which to directly manipulate the NADH/NAD$^+$ ratio in a tissue-specific manner and with subcellular resolution. Dietary supplements that raise hepatic NAD$^+$ levels in vivo do so indirectly and across a broad range of tissues[8], and classic redox tools such as methylene blue cannot be targeted to specific tissues and can affect multiple redox cofactors. Compounding this challenge, NADH and NAD$^+$ are compartmentalized within cells and organelles, with levels that differ by orders of magnitude, and exist in both free and protein-bound forms[10].

Recently, a genetic tool, a bacterial NADH oxidase from *Lb*NOX, was introduced that overcomes these limitations[1]. *Lb*NOX couples the oxidation of NADH to NAD$^+$ with the reduction of oxygen to water (Fig. 1a). It has high catalytic activity and can be genetically targeted to different subcellular compartments.

Here we used *Lb*NOX with metabolomics to characterize the biochemical consequences of lowering the hepatic free cytosolic NADH/NAD$^+$ ratio in vivo. Among the top changing metabolites is α-hydroxybutyrate (αHB), the elevation of which has previously been identified as a biomarker of early insulin resistance[3], impaired glucose tolerance[2], diabetes risk[11] and mitochondrial disorders[4]. The strongest metabolite quantitative trait locus for αHB is in the gene *GCKR*, which encodes glucokinase regulatory protein (GKRP), an inhibitor of hepatic glucokinase, and is among the most—if not the most—pleiotropic genome-wide association study (GWAS) loci, with over 100 studies linking common *GCKR* variation to over 130 metabolic traits and diseases (Supplementary Table 1). We provide evidence that many of these traits—including circulating αHB, triglyceride and fibroblast growth factor 21 (FGF21) levels and glucose tolerance—can lie downstream of the hepatic cytosolic NADH/NAD$^+$ ratio.

*Lb*NOX lowers NADH/NAD$^+$ in hepatocytes

We used adenovirus to express *Lb*NOX, mitochondrial-targeted *Lb*NOX (mito*Lb*NOX) or GFP in primary mouse hepatocytes (Extended

---

[1]Howard Hughes Medical Institute and Department of Molecular Biology, Massachusetts General Hospital, Boston, MA, USA. [2]Liver Center, Division of Gastroenterology, Massachusetts General Hospital, Boston, MA, USA. [3]Broad Institute, Cambridge, MA, USA. [4]Departments of Pediatrics and Genetics, Harvard Medical School, Boston, MA, USA. [5]Division of Endocrinology and Center for Basic and Translational Obesity Research, Boston Children's Hospital, Boston, MA, USA. [6]Department of Pathology, Massachusetts General Hospital, Boston, MA, USA. [7]Department of Neurobiology, Harvard Medical School, Boston, MA, USA. [8]Program in Molecular Medicine, University of Massachusetts Medical School, Worcester, MA, USA. [9]Division of Endocrinology, Metabolism, and Diabetes, Department of Medicine, University of Massachusetts Medical School, Worcester, MA, USA. [10]Department of Systems Biology, Harvard Medical School, Boston, MA, USA. ✉e-mail: vamsi@hms.harvard.edu

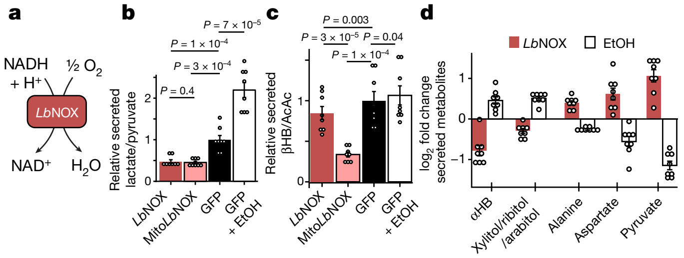


Fig. 1 | *Lb*NOX alters compartment-specific free NADH/NAD$^+$ ratio in hepatocytes. a, Chemical reaction catalysed by *Lb*NOX. **b**, **c**, The effect of *Lb*NOX, mito*Lb*NOX or ethanol (EtOH) on primary hepatocyte-secreted lactate/pyruvate (**b**) and hepatocyte-secreted βHB/AcAc (**c**) ratios. **d**, Relative changes in the abundance of secreted redox concordant metabolites. Data are mean ± s.e.m. from $n = 8$ independent hepatocyte isolations. Nominal *P* values were determined using paired, two-sided Student's *t*-tests between hepatocyte isolation.

Data Fig. 1a). As *Lb*NOX consumes oxygen, its activity in cells can be detected as non-mitochondrial oxygen consumption. Expression of *Lb*NOX or mito*Lb*NOX trends towards or increases, respectively, mitochondria-independent oxygen consumption (Extended Data Fig. 1b). The higher oxygen consumption of mito*Lb*NOX than of *Lb*NOX, which has previously been reported[1], is consistent with a higher free NADH concentration in mitochondria[10].

To assess the subcellular metabolic activity of *Lb*NOX, we measured secreted lactate/pyruvate and β-hydroxybutyrate (βHB)/acetoacetate (AcAc) ratios, which are metabolite ratios reflective of the free cytosolic and mitochondrial NADH/NAD$^+$ ratios, respectively[10]. *Lb*NOX markedly lowered the lactate/pyruvate ratio, whereas mito*Lb*NOX lowered both the βHB/AcAc and the lactate/pyruvate ratios (Fig. 1b, c). By contrast, ethanol, which generates cytosolic NADH through oxidation of ethanol via alcohol dehydrogenase, increased the lactate/pyruvate ratio. We also used the fluorescent reporter Peredox[12,13] to measure the free cytosolic NADH/NAD$^+$ ratio in hepatocytes exposed to different concentrations of ethanol (Extended Data Fig. 1c), with *Lb*NOX shifting the dose–response curve to the right. Thus, the combination of *Lb*NOX and ethanol allows selective perturbation of the free cytosolic NADH/NAD$^+$ ratio in hepatocytes.

Total cellular measurements of NAD$^+$, NADH, the NADH/NAD$^+$ ratio or total pool sizes (Extended Data Fig. 1d–g) are unchanged with *Lb*NOX expression. As total cellular measurements of NADH and NAD$^+$ integrate both free and bound levels originating from all cellular compartments, measurements with the lactate/pyruvate ratio or Peredox are more sensitive to changes in the free cytosolic NADH/NAD$^+$ ratio than total cellular measurements.

A screen for redox-sensitive metabolites

We next designed a screen for metabolites sensitive to the free cytosolic NADH/NAD$^+$ ratio. Primary hepatocytes were transduced with *Lb*NOX or with GFP ± ethanol, and the media were then profiled using liquid chromatography coupled to mass spectrometry. We focused on five metabolites with changes that were 'redox concordant', defined as metabolite levels that significantly increased with *Lb*NOX and decreased with ethanol, or vice versa (Fig. 1d, Supplementary Table 2). αHB was among the most sensitive and became the focus of further investigations given its previous association with several metabolic conditions.

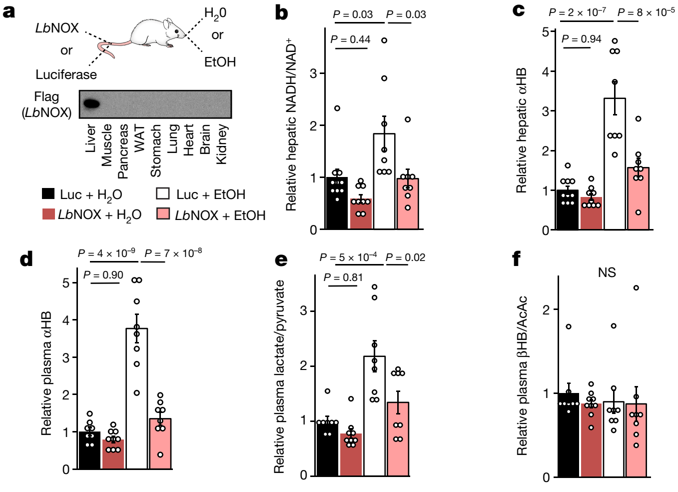


Fig. 2 | In vivo manipulation of the hepatic cytosolic NADH/NAD$^+$ ratio alters hepatic and circulating αHB levels. a, Experimental outline. WAT, white adipose tissue. **b**–**f**, The effects of the combination of hepatic *Lb*NOX or luciferase (Luc) expression and alcohol on the hepatic NADH/NAD$^+$ ratio (**b**), hepatic αHB levels (**c**), plasma αHB levels (**d**), the plasma lactate/pyruvate ratio (**e**) and the plasma βHB/AcAc ratio (**f**). Data are mean ± s.e.m. from $n = 8$–10 mice per group. *P* values were determined using one-way analysis of variance (ANOVA) with post-hoc Tukey's honestly significant difference (HSD) test. Not significant (NS) is $P = 0.90$.

The effects of *Lb*NOX and NR are distinct

NR supplementation has emerged as a robust means of increasing total NAD$^+$ pool sizes in vitro and in vivo[8,14], but its effect on cellular compartment-specific free NADH/NAD$^+$ ratios is unclear. We therefore performed a direct comparison of the metabolic effects of NR versus *Lb*NOX in hepatocytes. In contrast to *Lb*NOX, NR markedly boosted total NAD$^+$ levels and the total NAD(H) pool size (Extended Data Fig. 2a–c), yet did not lower compartment-specific free NADH/NAD$^+$ ratios as measured by lactate/pyruvate or βHB/AcAc ratios, nor did it lower αHB levels (Extended Data Fig. 2d–f). We investigated whether NR supplementation could rescue the pyruvate auxotrophy that is classically observed in the context of electron transport chain inhibition and is known to be due to an elevated cytosolic NADH/NAD$^+$ ratio[1,15]. In the presence of piericidin A (an inhibitor of mitochondrial complex I), either pyruvate or *Lb*NOX expression could rescue cellular proliferation, but NR supplementation could not (Extended Data Fig. 2g), despite boosting total cellular NAD$^+$ levels (Extended Data Fig. 2h). This illustrates both how measurements of total cellular NAD(H) levels provide little information with respect to compartment-specific free NAD(H) and suggests that the increased cellular NAD$^+$ from NR may contribute to the free NADH/NAD$^+$ pools in a redox-neutral manner.

Hepatic NADH/NAD$^+$ alters αHB in vivo

To extend these findings in vivo, we evaluated the metabolic effects in mice that were tail-vein-injected with *Lb*NOX or luciferase control adenovirus and gavaged with ethanol or water (Fig. 2a). We observed that alcohol increased the hepatic NADH/NAD$^+$ ratio and αHB levels, which were largely prevented with *Lb*NOX expression (Fig. 2b, c), and an identical pattern between plasma αHB levels and the lactate/pyruvate ratio (Fig. 2d, e), but not the βHB/AcAc ratio (Fig. 2f), demonstrating that αHB is a biomarker of the hepatic free cytosolic NADH/NAD$^+$ ratio in vivo. To complement these findings in another model, we examined the *Ndufs4*-knockout mouse model of mitochondrial disease due to deficiency in complex I[16], the primary enzyme responsible for oxidation of NADH to NAD$^+$ in most cell types. The *Ndufs4*-knockout mice had significantly higher plasma αHB levels and hepatic NADH/NAD$^+$ ratios, consistent with previous findings[17] (Extended Data Fig. 3).

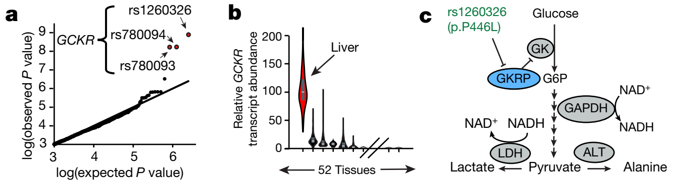


Fig. 3 | A common *GCKR* variant is associated with plasma αHB levels in humans. a, *Q*–*Q* plot of variants associated with plasma αHB levels, from ref. [5]. **b**, Relative distribution of *GCKR* expression in humans from the Genotype-Tissue Expression (GTEx) Project[44]. The violin plots show the median, and the 25th and 75th percentile expression. **c**, Pathway diagram that depicts the role of GKRP (encoded by *GCKR*) and its polymorphism. ALT, alanine transaminase; G6P, glucose 6-phosphate; GK, glucokinase; LDH, lactate dehydrogenase.

A *GCKR* variant raises hepatic NADH/NAD$^+$

We next hypothesized whether we could identify candidate human genetic determinants of the hepatic NADH/NAD$^+$ ratio via genetic association studies of plasma αHB levels. Data from a metabolite quantitative trait locus study of the Framingham Heart Study reveal that the top three single-nucleotide polymorphisms associated with αHB levels—rs1260326, rs780094 and rs780093[5] (Fig. 3a)—are in the *GCKR* locus, a liver-specific gene (Fig. 3b). The three single-nucleotide polymorphisms are in tight linkage disequilibrium and the 'risk' haplotype, which approaches 50% prevalence in some populations[18] (Extended Data Fig. 4a), has been associated with a notably large number of disease states and metabolic traits, spanning cardiometabolic risk factors such as circulating cholesterol and triglyceride levels[19,20], plasma levels of small circulating metabolites[5], behavioural traits such as alcohol and coffee consumption[21,22], important mediators of metabolic signalling such as leptin[23] and FGF21[24], as well as diseases such as diabetes[25] and fatty liver disease[26] (Supplementary Table 1).

GKRP (encoded by *GCKR*) sequesters hepatic glucokinase in the liver during fasting and releases it during feeding. It is thought to help to prevent a futile metabolic cycle during gluconeogenesis with glucose 6-phosphatase and provide a rapid mechanism to increase hepatic glucose metabolism during feeding[27] (Fig. 3c). rs1260326 encodes a P446L missense mutation in *GCKR*, which inhibits the activity of GKRP and results in higher glucokinase activity and hepatic glycolytic flux[28], and is the likely causal variant that underlies traits associated with the risk *GCKR* haplotype.

To our knowledge, no previous study has invoked the hepatic NADH/NAD$^+$ ratio as an effector of *GCKR* variants. To investigate this possibility, we identified plasma metabolites from the experiment shown in Fig. 2 that were sensitive to the hepatic NADH/NAD$^+$ ratio (Methods), and then calculated an enrichment score for these metabolites across all metabolite quantitative trait loci reported in a previous study[5]. This analysis revealed *GCKR* as among the top loci for redox-sensitive metabolites (Extended Data Fig. 4b).

We then measured the metabolic consequences of overexpression of either GKRP-446P or GKRP-446L (the rs1260326 variant) in primary hepatocytes (Extended Data Fig. 4c). GKRP-446P overexpression decreased the lactate/pyruvate ratio and αHB levels, but this effect was less with the 446L variant, despite comparable levels of overexpression. This indicates that *GCKR* directly influences the free cytosolic NADH/NAD$^+$ ratio, and that this effect is blunted for the 446L variant. In addition, the directions of change in lactate, pyruvate, alanine and αHB levels were all consistent with changes observed at the *GCKR* locus in previously reported GWAS (Supplementary Table 1). Glucose is a fundamental source of reducing equivalents, and inhibition of glycolytic flux caused by GKRP is diminished with the 446L variant[28]. Thus, we speculate that the effect of *GCKR* on decreasing the free hepatic NADH/NAD$^+$ ratio is most probably through modulation of glycolytic flux and/or its fate.

*Lb*NOX alters insulin resistance in vivo

As *GCKR* variants are associated with impaired glucose tolerance[29] and circulating αHB levels are a biomarker of impaired glucose tolerance and insulin resistance[3], we investigated the effects of *Lb*NOX expression in the HFD mouse model of insulin resistance and dysglycaemia.

At baseline, HFD mice had elevated circulating levels of αHB compared to chow-fed diet mice, which were significantly lowered with *Lb*NOX (Fig. 4a). Hepatic *Lb*NOX improved glucose tolerance in HFD mice (Fig. 4b) and hepatic insulin resistance during hyperinsulinaemic–euglycaemic clamps (Fig. 4c, d), although not any other measured parameters including body weight or insulin-stimulated glucose uptake in peripheral tissues (Extended Data Fig. 5, Supplementary Table 3). These data indicate a causal link between the free cytosolic hepatic NADH/NAD$^+$ ratio and hepatic insulin resistance in our system, which potentially underlies the association of αHB with insulin resistance in humans[3].

Most models of hepatic insulin resistance invoke a disruption of hepatic insulin signalling at or upstream of AKT via several proposed mediators[30–33], such as diacyl glycerols or ceramides, leading to downstream alterations in forkhead box protein O1 (FOXO1) activity and transcription of gluconeogenic genes. However, we could find no difference in hepatic diacyl glycerol or ceramide content with *Lb*NOX expression (Extended Data Fig. 6), nor acute insulin-mediated AKT phosphorylation (Extended Data Fig. 7a–c), FOXO1 target transcript abundance (Extended Data Fig. 7d), or AKT phosphorylation or FOXO1 targets at the end of the hyperinsulinaemic–euglycaemic clamp (Extended Data Fig. 7e–g). Total NAD$^+$, NADPH and NADP$^+$ pool sizes were also unchanged (Extended Data Fig. 8).

Measurements of the gluconeogenic intermediates that sequentially link the conversion of pyruvate to glucose, with the exception of glyceraldehyde 3-phosphate and oxaloacetate, which we could not reliably measure, suggest a control point at either glyceraldehyde-3-phosphate dehydrogenase (GAPDH) or triosephosphate isomerase (Extended Data Fig. 7h, top, Extended Data Fig. 7i), which is most probably at GAPDH as it requires NADH to reduce bisphosphoglycerate to glyceraldehyde 3-phosphate. This is further strengthened by an analysis of livers grouped into high or low lactate/pyruvate ratio, which even more clearly identifies the GAPDH or triosephosphate isomerase crossover point (Extended Data Fig. 7h, bottom).

NADH/NAD$^+$ can influence *GCKR*-linked traits

Having established that the free hepatic cytosolic NADH/NAD$^+$ ratio underlies the association between *GCKR* and αHB, we hypothesized that some of the many other traits linked to *GCKR* variation might similarly be mediated by the hepatic NADH/NAD$^+$ ratio (Fig. 5a).

To investigate this possibility, we used various analytic methods (Methods) to measure the plasma levels of 51 analytes associated with *GCKR* in GWAS using the experimental scheme shown in Fig. 2 (Fig. 5b). Of these, we found that 28 were sensitive to the hepatic cytosolic NADH/NAD$^+$ ratio, of which nearly all were in the same direction as their GWAS associations. Although some tested metabolites did not exhibit sensitivity to the hepatic NADH/NAD$^+$ ratio including plasma glucose levels (Fig. 5c), many did and notably included plasma triglyceride levels (Fig. 5d) (the most replicated GWAS association with the *GCKR* risk haplotype in over 30 studies; Supplementary Table 1), plasma serine levels (Fig. 5e) and circulating FGF21 levels (Fig. 5f).

Not all *GCKR*-linked traits that we could measure showed NADH/NAD$^+$ sensitivity, perhaps owing to technical reasons (for example, our acute experimental paradigm versus chronic human changes) or the fact that some *GCKR* associations may operate via changes in hepatic glycolytic flux that are non-redox-based.

This approach demonstrates the effectiveness of using *GCKR* GWAS associations as human genetic tools to uncover novel human

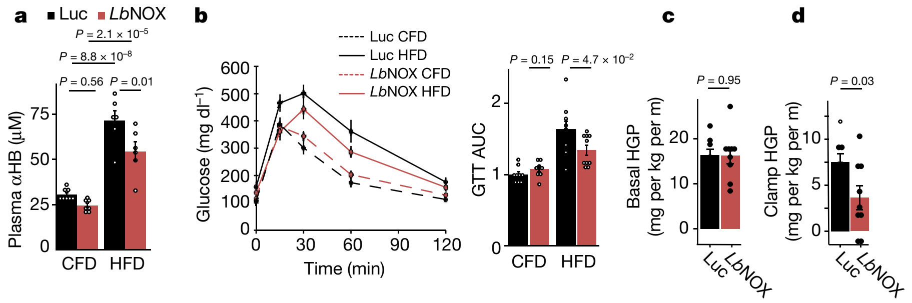


Fig. 4 | Direct oxidation of free hepatic cytosolic NADH improves glucose tolerance and hepatic insulin sensitivity in vivo. a, **b**, The effects of hepatic *Lb*NOX expression on plasma αHB levels (**a**) and glucose tolerance (**b**) in chow-fed (CFD) or HFD mice. AUC, area under the curve; GTT, glucose tolerance test. **c**, **d**, The effect of hepatic *Lb*NOX in HFD-fed mice during hyperinsulinaemic–euglycaemic clamp on basal hepatic glucose production (HGP) (**c**) and clamp HGP (**d**). *P* values were determined using unpaired, two-sided Student's *t*-test (**b**–**d**) or one-way ANOVA with post-hoc Tukey's HSD test (**a**). Data are mean ± s.e.m. from *n* = 6–8 mice (**a**), 9–10 (**b**) and 8–9 total (**c**).

NADH/NAD$^+$ biology, and further supports the concept that the hepatic redox state is an effector of *GCKR* to influence metabolic traits.

Discussion

With the use of *Lb*NOX, we have been able to change the free NADH/NAD$^+$ ratio in a specific cellular compartment (the cytoplasm) and in a specific tissue (the liver) in vivo. In doing so, we have identified αHB as a circulating biomarker of the free hepatic cytosolic NADH/NAD$^+$ ratio, which is in turn influenced by common human genetic variation in *GCKR*. This has allowed us to identify excess NADH electrons, or reductive stress, as an effector of many metabolic traits linked to *GCKR* through GWAS (Fig. 5g).

αHB is a relatively obscure metabolite, the only known production of which in humans is through the reduction of α-ketobutyrate via lactate dehydrogenase in the cytosol (Extended Data Fig. 9). An elevation in the cytosolic NADH/NAD$^+$ ratio would therefore promote αHB formation and probably underlies our observation that αHB is a biomarker of the cytosolic NADH/NAD$^+$ ratio. Although we cannot exclude the possibility of non-hepatic contributions to the circulating αHB pool, recent work suggests that only the liver and, interestingly, skin are significant sources[34]. We note that other sources of αKB could also potentially influence αHB levels.

Previous metabolomics studies have demonstrated that elevated plasma αHB levels are associated with dysglycaemia, and we show both that the level of αHB is elevated in a mouse model of insulin resistance and that direct oxidation of hepatic cytosolic NADH in this setting improves hepatic insulin resistance, possibly independent of canonical intrinsic hepatic insulin signalling[35,36]. Our data phenocopy the findings in humans that lower αHB levels are associated with improved glucose tolerance and insulin resistance but not basal blood glucose levels[2,3]. Our data also clarify causal relationships that underlie the mechanism by which *GCKR* segregates dysglycaemic traits, namely, how the *GCKR* variant that worsens glucose tolerance[29,37] improves basal glucose levels[25]: it is possible that the blood-glucose-lowering effect of the *GCKR* variant causally increases reductive stress, and this high NADH/NAD$^+$ ratio is causally related to hepatic insulin resistance and glucose tolerance. We also note recent work demonstrating that a *SLC16A11* haplotype, which increases the risk of type II diabetes, also increases the cytosolic NADH/NAD$^+$ ratio, and hypothesize that the associated reductive stress is causally related to its phenotype[38].

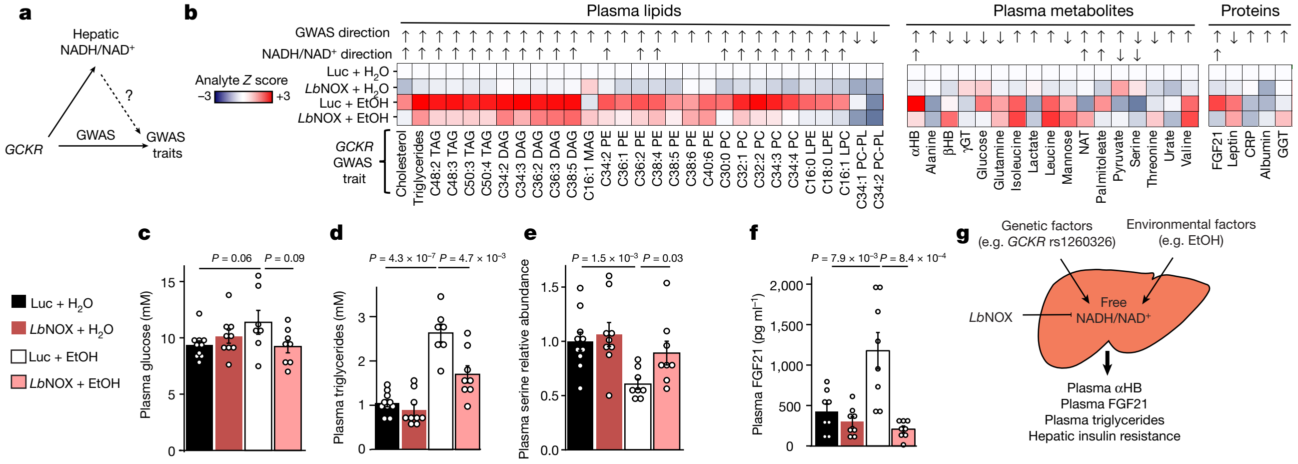


Fig. 5 | Many *GCKR*-associated metabolic traits lie downstream of hepatic NADH reductive stress. a, GWAS have linked many traits to *GCKR* variation. The current study raises the hypothesis that some of these are mediated by variation in the hepatic NADH/NAD$^+$ ratio. **b**, Fifty-one such traits that we could measure using the EtOH/*Lb*NOX in vivo system in Fig. 2a are shown, with the analyte *Z* score for each condition relative to Luc + H_2O, whether the measured analyte fulfilled our criteria for 'NADH/NAD$^+$ sensitivity' (Methods), and if so, its direction, along with the observed direction of effect, from the P446L risk haplotype in published studies (Supplementary Table 1). CRP, C-reactive protein; DAG, diacylglycerol; GGT, γ-glutamyltransferase; γGT, γ-glutamylthreonine; LPC, lysophosphatidylcholine; LPE, lysophosphatidylethanolamine; MAG, monoacylglycerol; NAT, *N*-acetyltryptophan; PC, phosphatidylcholine; PC-PL, phosphatidylcholine plasmalogen; PE, phosphatidylethanolamine; TAG, triacylglycerol. **c**–**f**, Selected data are shown for plasma glucose levels (**c**), total plasma triglyceride levels (**d**), plasma serine levels (**e**) and plasma FGF21 levels (**f**). **g**, Proposed model, in which circulating αHB is a biomarker of an elevated free hepatic cytosolic NADH/NAD$^+$ ratio, which is a latent metabolic parameter that is influenced by genetic and environmental factors and serves as an effector of different metabolic traits. Data are mean ± s.e.m. from *n* = 7–10 mice per group. Nominal *P* values were determined using two-sided Student's *t*-tests for the 'NADH/NAD$^+$' sensitivity calculations.

Motivated by the association between circulating levels of αHB and *GCKR* genetic variation[5], we are able to demonstrate that common human genetic variation in *GCKR* directly influences hepatic reductive stress, and that many traits linked to *GCKR* variants are influenced by reductive stress in a mouse model (Fig. 5b).

Our work has uncovered two notable traits that have previously been associated with *GCKR* variation, but with a link to hepatic NADH redox metabolism that had not previously been reported: circulating triglyceride levels (Fig. 5d) and plasma FGF21 levels (Fig. 5f). Plasma triglyceride levels, a cardiometabolic risk factor, are the most replicated *GCKR* GWAS association (Supplementary Table 1), nominating hepatic NADH reductive stress as a latent risk factor of cardiometabolic disease. Reductive stress also probably contributes to the increase in circulating triglyceride levels observed with heavy alcohol consumption[39], and may be causally related to hepatic fat accumulation seen with heavy alcohol use, as well as hepatic fat more generally given the association of *GCKR* with hepatic fat content in non-alcoholic fatty liver disease[26]. FGF21 is a hepatokine that has many metabolic actions, and although a redox basis for its secretion has never before been proposed, previous studies have shown that alcohol consumption causes its release[40].

We speculate that the simultaneous release of triglycerides and FGF21 by the liver in response to NADH reductive stress may represent an adaptive, dissipative programme. Triglycerides are energy-rich lipids that can be packaged and exported to peripheral tissues. FGF21 is known to boost metabolic rates through increased uncoupling[41] or energy expenditure[42] and specifically promotes the clearance of circulating triglycerides by promoting their uptake into fat[43]. The release of triglycerides and FGF21 in response to elevations in the NADH/NAD$^+$ ratio may therefore represent a whole-body reductive stress response, the logic of which is to transfer reducing equivalents from the liver into peripheral tissues for storage or for catabolism. Future challenges lie in testing this model and elucidating the molecular mechanisms that link high NADH/NAD$^+$ ratios in the cytosol to elevations in circulating FGF21 and triglyceride levels.

Through the use of *Lb*NOX—a genetic tool that enables causal studies of NADH/NAD$^+$ metabolism—our work strongly supports the concept that the alterations in αHB levels seen in insulin resistance, common *GCKR* variants and mitochondrial disorders reflect an excess of hepatic reducing equivalents, and that this hepatic reductive stress can be a causal determinant for many metabolic traits and diseases (Fig. 5g). Future efforts will be required to directly target reductive stress for therapeutic benefit, as well as define adaptive and maladaptive components of the reductive stress response.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2337-2.

1. Titov, D. V. et al. Complementation of mitochondrial electron transport chain by manipulation of the NAD$^+$/NADH ratio. *Science* **352**, 231–235 (2016).
2. Cobb, J. et al. α-Hydroxybutyric acid is a selective metabolite biomarker of impaired glucose tolerance. *Diabetes Care* **39**, 988–995 (2016).
3. Gall, W. E. et al. α-Hydroxybutyrate is an early biomarker of insulin resistance and glucose intolerance in a nondiabetic population. *PLoS ONE* **5**, e10883 (2010).
4. Thompson Legault, J. et al. A metabolic signature of mitochondrial dysfunction revealed through a monogenic form of Leigh syndrome. *Cell Rep.* **13**, 981–989 (2015).
5. Rhee, E. P. et al. A genome-wide association study of the human metabolome in a community-based cohort. *Cell Metab.* **18**, 130–143 (2013).
6. Goodman, R. P., Calvo, S. E. & Mootha, V. K. Spatiotemporal compartmentalization of hepatic NADH and NADPH metabolism. *J. Biol. Chem.* **293**, 7508–7516 (2018).
7. Trammell, S. A. J. et al. Nicotinamide riboside opposes type 2 diabetes and neuropathy in mice. *Sci. Rep.* **6**, 26933 (2016).
8. Cantó, C. et al. The NAD$^+$ precursor nicotinamide riboside enhances oxidative metabolism and protects against high-fat diet-induced obesity. *Cell Metab.* **15**, 838–847 (2012).
9. Madiraju, A. K. et al. Metformin suppresses gluconeogenesis by inhibiting mitochondrial glycerophosphate dehydrogenase. *Nature* **510**, 542–546 (2014).
10. Krebs, H. A. The redox state of nicotinamide adenine dinucleotide in the cytoplasm and mitochondria of rat liver. *Adv. Enzyme Regul.* **5**, 409–434 (1967).
11. Ferrannini, E. et al. Early metabolic markers of the development of dysglycemia and type 2 diabetes and their physiological significance. *Diabetes* **62**, 1730–1737 (2013).
12. Hung, Y. P., Albeck, J. G., Tantama, M. & Yellen, G. Imaging cytosolic NADH-NAD$^+$ redox state with a genetically encoded fluorescent biosensor. *Cell Metab.* **14**, 545–554 (2011).
13. Masia, R. et al. Live cell imaging of cytosolic NADH/NAD+ ratio in hepatocytes and liver slices. *Am. J. Physiol. Gastrointest. Liver Physiol.* **314**, G97–G108 (2018).
14. Trammell, S. A. J. et al. Nicotinamide riboside is uniquely and orally bioavailable in mice and humans. *Nat. Commun.* **7**, 12948 (2016).
15. King, M. P. & Attardi, G. Human cells lacking mtDNA: repopulation with exogenous mitochondria by complementation. *Science* **246**, 500–503 (1989).
16. Kruse, S. E. et al. Mice with mitochondrial complex I deficiency develop a fatal encephalomyopathy. *Cell Metab.* **7**, 312–320 (2008).
17. Jain, I. H. et al. Hypoxia as a therapy for mitochondrial disease. *Science* **352**, 54–61 (2016).
18. Machiela, M. J. & Chanock, S. J. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* **31**, 3555–3557 (2015).
19. Teslovich, T. M. et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707–713 (2010).
20. Diabetes Genetics Initiative of Broad Institute of Harvard and MIT et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* **316**, 1331–1336 (2007).
21. Cornelis, M. C. et al. Genome-wide meta-analysis identifies six novel loci associated with habitual coffee consumption. *Mol. Psychiatry* **20**, 647–656 (2015).
22. Schumann, G. et al. *KLB* is associated with alcohol drinking, and its gene product β-Klotho is necessary for FGF21 regulation of alcohol preference. *Proc. Natl Acad. Sci. USA* **113**, 14372–14377 (2016).
23. Kilpeläinen, T. O. et al. Genome-wide meta-analysis uncovers novel loci influencing circulating leptin levels. *Nat. Commun.* **7**, 10494 (2016).
24. Cheung, C. Y. Y. et al. An exome-chip association analysis in Chinese subjects reveals a functional missense variant of *GCKR* that regulates FGF21 levels. *Diabetes* **66**, 1723–1728 (2017).
25. Orho-Melander, M. et al. Common missense variant in the glucokinase regulatory protein gene is associated with increased plasma triglyceride and C-reactive protein but lower fasting glucose concentrations. *Diabetes* **57**, 3112–3121 (2008).
26. Speliotes, E. K. et al. Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. *PLoS Genet.* **7**, e1001324 (2011).
27. Raimondo, A., Rees, M. G. & Gloyn, A. L. Glucokinase regulatory protein: complexity at the crossroads of triglyceride and glucose metabolism. *Curr. Opin. Lipidol.* **26**, 88–95 (2015).
28. Beer, N. L. et al. The P446L variant in GCKR associated with fasting plasma glucose and triglyceride levels exerts its effect through increased glucokinase activity in liver. *Hum. Mol. Genet.* **18**, 4081–4088 (2009).
29. Dupuis, J. et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat. Genet.* **42**, 105–116 (2010).
30. Holland, W. L. & Summers, S. A. Sphingolipids, insulin resistance, and metabolic disease: new insights from in vivo manipulation of sphingolipid metabolism. *Endocr. Rev.* **29**, 381–402 (2008).
31. Perry, R. J., Samuel, V. T., Petersen, K. F. & Shulman, G. I. The role of hepatic lipids in hepatic insulin resistance and type 2 diabetes. *Nature* **510**, 84–91 (2014).
32. Nakatani, Y. et al. Involvement of endoplasmic reticulum stress in insulin resistance and diabetes. *J. Biol. Chem.* **280**, 847–851 (2005).
33. Shoelson, S. E., Lee, J. & Goldfine, A. B. Inflammation and insulin resistance. *J. Clin. Invest.* **116**, 1793–1801 (2006).
34. Jang, C. et al. Metabolite exchange between mammalian organs quantified in pigs. *Cell Metab.* **30**, 594–606.e3 (2019).
35. Titchenell, P. M., Chu, Q., Monks, B. R. & Birnbaum, M. J. Hepatic insulin signalling is dispensable for suppression of glucose output by insulin in vivo. *Nat. Commun.* **6**, 7078 (2015).
36. Ader, M. & Bergman, R. N. Peripheral effects of insulin dominate suppression of fasting hepatic glucose production. *Am. J. Physiol.* **258**, E1020–E1032 (1990).
37. Saxena, R. et al. Genetic variation in *GIPR* influences the glucose and insulin responses to an oral glucose challenge. *Nat. Genet.* **42**, 142–148 (2010).
38. Kim, W. et al. Polyunsaturated fatty acid desaturation is a mechanism for glycolytic NAD$^+$ recycling. *Cell Metab.* **29**, 856–870.e7 (2019).
39. Baraona, E. & Lieber, C. S. Effects of ethanol on lipid metabolism. *J. Lipid Res.* **20**, 289–315 (1979).
40. Desai, B. N. et al. Fibroblast growth factor 21 (FGF21) is robustly induced by ethanol and has a protective role in ethanol associated liver injury. *Mol. Metab.* **6**, 1395–1406 (2017).
41. Hondares, E. et al. Hepatic FGF21 expression is induced at birth via PPARα in response to milk intake and contributes to thermogenic activation of neonatal brown fat. *Cell Metab.* **11**, 206–212 (2010).
42. Xu, J. et al. Fibroblast growth factor 21 reverses hepatic steatosis, increases energy expenditure, and improves insulin sensitivity in diet-induced obese mice. *Diabetes* **58**, 250–259 (2009).
43. Schlein, C. et al. FGF21 lowers plasma triglycerides by accelerating lipoprotein catabolism in white and brown adipose tissues. *Cell Metab.* **23**, 441–453 (2016).
44. Lonsdale, J. et al. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013).

Methods

Primary hepatocyte experiments

Primary hepatocytes were freshly isolated by perfusion and enzymatic digestion (5401119001, Roche) of livers from male C57BL/6J mice, aged 12–16 weeks, and plated on six-well collagen-coated plates (A1142801, Life Technologies) at a density of 4×10^5 cells per well in DMEM medium (11995-065, Gibco) supplemented with 10% FBS (F2442, Sigma-Aldrich), 200 U/ml penicillin–streptomycin. Cells were transduced with adenovirus immediately after isolation, and all subsequent experiments were performed 24 h after isolation. Before the metabolomics experiments, transduced hepatocytes were placed in serum-free medium (A1443001, Gibco; DMEM supplemented with 5.5 mM glucose), and medium samples were collected 2 h (Fig. 1, Extended Data Fig. 1) or 4 h later (Extended Data Figs. 2, 4c). For experiments with NR, NR chloride (Tru Niagen) in PBS was added to a final concentration of 500 µM at the time of hepatocyte seeding, and included at a concentration of 500 µM in the serum-free medium for use in the metabolomics experiments. For primary hepatocyte experiments, see Fig. 1b–d, Extended Data Figs. 1, 2, 4c.

Rescue of piericidin-induced inhibition of cell proliferation

Ten thousand HeLa Tet3G *Lb*NOX cells (previously described in ref. [1]) were seeded in 24-well plates in DMEM 11995-065 medium supplemented with 10% FBS (F2442, Sigma-Aldrich), 200 U/ml penicillin–streptomycin. Twenty-four hours after seeding, either water or doxycycline was added to a final concentration of 300 ng/ml in each well. Twenty-four hours later, media were exchanged to DMEM without pyruvate (D9802, US Biological), supplemented with 10% dialysed FBS (26400-044, Life Technologies), 200 µM uridine, ± 1 µM piericidin A, ± 1 mM pyruvate, ± 500 µM NR and ± 300 ng/ml doxycycline. Cells were counted 4 days later. For the piericidin-induced inhibition experiment, see Extended Data Fig. 2g, h.

Adenovirus production, amplification and titring

All adenoviruses were type 5 (dE1/dE3), with gene expression driven by a CMV promoter. Seed GFP and luciferase control adenoviruses were purchased from Vector Biolabs and Signagen, respectively. *Lb*NOX and mito*Lb*NOX adenoviruses were generated using the ViraPower Adenoviral Expression Kit (K493000, Thermo Fisher) after subcloning from pUC57-*Lb*NOX and pUC57-mito*Lb*NOX (Addgene Plasmids 75275 and 75285, respectively). Mouse *GCKR* adenovirus (GKRP-446P (ref. BC012412) and GKRP-446L) were purchased from Vector Biolabs. All viruses were amplified in 293A cell lines (from the American Type Culture Collection) and titred using the Adeno-X Rapid Titer Kit (632250, Takara Bio), co-titred with aliquots of a standard virus of known titre as a control. High-titre virus for in vivo expression was purified using the Adenovirus Purification Kit (003054, ViraPur).

Primary hepatocyte oxygen consumption

Hepatocyte respiration was measured 24 h after isolation and transduction in a Seahorse Bioscience XF24-3 Analyzer. Before seeding, Seahorse XF24 cell plates were coated with 30 µl of 100 µg/ml collagen I from rat tail (ALX-522-435, Enzo). Immediately after hepatocyte isolation, fresh hepatocytes were added to cell plates at a density of 5,000 cells per well in 250 µl of hepatocyte isolation medium (described above), in addition to the specified adenovirus, and incubated overnight. Twenty-four hours after plating, media were replaced with 450 µl of assay medium (DMEM A14430-1, Gibco; 10% FBS, 200 U/ml penicillin–streptomycin, 5.5 mM HEPES, 1 nM insulin, 1 nM dexamethasone, 4.5 g/l glucose), and oxygen consumption was measured. Each measurement was performed after a 4-min mix period, a 30-s pause and a 2-min wait period. Electron transport chain inhibitors antimycin A and rotenone were injected to a final concentration of 2 µM each. For the primary hepatocyte oxygen consumption experiment, see Extended Data Fig. 1b.

Imaging Peredox in mouse hepatocytes

Freshly isolated hepatocytes were plated on fibronectin-coated glass coverslips in a six-well plate and incubated for 4 h in hepatocyte medium with adenovirus. The transduction mixture was removed, and hepatocytes were transfected with 1.0 µg Peredox-mCherry in pcDNA3.1 using Effectene transfection reagent (Qiagen) in fresh medium. After 12–14 h, the transfection mixture was removed and fresh medium was added. Experiments were carried out 24–32 h afterwards (36–48 h after hepatocyte isolation). Wide-field epifluorescence experiments were performed in a diamond-shaped solution chamber mounted on the headstage of an inverted microscope (Nikon) under continuous perfusion (1.2 ml/min flow rate) heated at 37 °C. Glass coverslips containing plated hepatocytes were cracked with a diamond-tip pen and shards were placed in the recording chamber. The bath solution was 140 mM NaCl, 10 mM glucose, 10 mM HEPES, 5 mM KCl, 2 mM $CaCl_2$, 1 mM $MgCl_2$, and was pH-adjusted to 7.4 with 1 M NaOH. Ethanol was added as indicated. Glass shards were preincubated in bath solution at 37 °C for 10 min before placing in the recording chamber. Emitted light was collected with an Andor Revolution DSD spinning disk unit (Andor) using a ×20/0.75 NA objective illuminated with a LED light source (Lumencor). The green and red fluorophores of Peredox were excited using 405/10-nm and 578/16-nm band-pass filters, emission was collected through 525/50-nm and 629/56-nm band-pass filters, and excitation and emission light were separated with 490-nm and 590-nm short-pass dichroics, respectively. Images were acquired with iQ (Andor) every 15 s at 50-ms exposure and 4 × 4 binning.

Fluorescence intensity was quantified with ImageJ as the mean over a region of interest drawn around each analysed hepatocyte. The Peredox signal (green/red) was expressed as a per cent value in which 0 is the 'floor' (obtained in 10 mM pyruvate) and 100 is the 'ceiling' (obtained in 10 mM ethanol). For dose–response curves, a Hill function was fit to the data using the Solver function in Excel. For the Peredox experiments, see Extended Data Fig. 1c.

Analyte measurements

Mouse blood glucose was measured with an AlphaTRAK2 glucose monitor (Zoetis), plasma glucose with Glucose-SL reagent (235-60, Sekisui), triglycerides with Triglyceride-SL reagent (236-99, Sekisui), leptin with a mouse leptin ELISA (ADI-900-019A, Enzo), FGF21 with a mouse FGF21 ELISA (EZRMFGF21-26K, Sigma Millipore), GGT with a mouse GGT ELISA (OKEH03351, Aviva Systems Biology), CRP with a mouse CRP ELISA (RAB1121, Sigma) and albumin with a BCG Albumin Assay (MAK124, Sigma). Whole-cell or whole-tissue NADH/NAD^+ for Extended Data Figs. 1d–g, 2a–c, h were done with NAD^+/NADH-Glo Assay (G9071, Promega).

Analytes not listed above were measured with liquid chromatography coupled to mass spectrometry (LC–MS) or gas chromatography (GC)–MS as described below.

Animal experiments

All animal experiments in this paper were approved by the Massachusetts General Hospital or University of Massachusetts Institutional Animal Care and use Committee, and all relevant ethical regulations were followed.

Male C57BL/6J mice aged 12–16 weeks were purchased from The Jackson Laboratory, and were administered chow (Prolab Isorpo RMH 3000 5p75) and water ad libitum. For diet-induced obesity (DIO) mice, male C57BL/6J mice aged 12–16 weeks were fed a HFD (D12492, Research Diets) or were purchased from The Jackson Laboratory and maintained on their diet before experiments.

For adenoviral experiments, $2–4 \times 10^9$ plaque-forming units of adenovirus were given via tail-vein injection, and experiments were performed 4 days post-injection, after which time mice were killed. For each experiment, different conditions (that is, *Lb*NOX or luciferase) were randomly divided among cagemates. Investigators were not blinded to the identity of mice, and sample sizes were not prespecified.

For gavage experiments, an oral gavage of 3.5 g/kg or equivalent H_2O volume for controls was given at the start of the experiment, followed by a second gavage of half the initial dose 1 h later. Plasma and liver tissue were collected 6 h after the initial dose under isoflurane anaesthesia and immediately stored at −80 °C (plasma) or flash frozen in liquid nitrogen (liver) until further processing, as described below.

For glucose tolerance tests, overnight fasted mice were given an intraperitoneal injection of 2g/kg of 20% glucose in saline. Blood glucose levels were subsequently measured using an AlphaTRAK2 glucometer. For acute insulin experiments, DIO mice were fasted for 6 h and then injected with 2 U/kg insulin, with liver tissue collected 15 min later under isoflurane anaesthesia and flash frozen in liquid nitrogen before further analysis. For mouse experiments with *Ndufs4*-knockout (KO) mice, liver and plasma were collected from unfasted wild-type and KO littermates at age 48 days. For plasma αHB measurements in CFD versus HFD mice, as well as qPCR of FOXO1 transcriptional targets, mice were fasted for 6 h before measurements.

Hyperinsulinaemic–euglycaemic clamp

Hyperinsulinaemic–euglycaemic clamps were conducted as previously described[45]. In brief, survival surgery was performed at 5–6 days before clamp experiments to establish an indwelling catheter in the jugular vein, which was used to inject 2×10^9 plaque-forming units of viral vector 4 days before the clamp. On the day of the clamp experiment, mice were fasted overnight (about 17 h), and a 2-h hyperinsulinaemic–euglycaemic clamp was conducted in conscious mice with a primed and continuous infusion of human insulin (150 mU/kg body weight priming followed by 2.5 mU/kg/min; Humulin, Eli Lilly). To maintain euglycaemia, 20% glucose was infused at variable rates during clamps. Whole-body glucose turnover was assessed with a continuous infusion of [3-^{3}H] glucose (PerkinElmer), and 2-deoxy-D-[1-^{14}C]glucose (2-[^{14}C]DG) (PerkinElmer) was administered as a bolus (10 μCi) at 75 min after the start of clamps to measure insulin-stimulated glucose uptake in individual organs. At the end of the clamps, mice were anaesthetized and tissues were taken for biochemical analysis.

Glucose concentrations during clamps were analysed using 5–10 μl plasma by a glucose oxidase method on an Analox GM9 Analyser (Analox Instruments). Plasma concentrations of [3-^{3}H]glucose, 2-[^{14}C] DG and ^3H_2O were determined following deproteinization of plasma samples as previously described. For the determination of tissue 2-[^{14}C] DG-6-phosphate content, tissue samples were homogenized and the supernatants were subjected to an ion-exchange column to separate 2-[^{14}C]DG-6-phosphate from 2-[^{14}C]DG.

Rates of basal HGP and insulin-stimulated whole-body glucose turnover were determined as previously described[45]. The insulin-stimulated rate of HGP was determined by subtracting the glucose infusion rate from whole-body glucose turnover. Whole-body glycolysis and glycogen plus lipid synthesis from glucose were calculated as previously described[45]. Insulin-stimulated glucose uptake in individual tissues was assessed by determining the tissue (for example, skeletal muscle) content of 2-[^{14}C]DG-6-phosphate and the plasma 2-[^{14}C]DG profile. For the hyperinsulinaemic–euglycaemic clamp experiments, see Fig. 4c, d, Extended Data Fig. 5a–k).

Metabolomics experiments

For the media and plasma metabolite measurements (Fig. 1b–d, 2d–f, 4a, 5b, d, e, Extended Figs. 2d–f, 3a, 4c): medium or plasma sample (30 μl) were mixed with 137 μl of ice-cold acetonitrile containing internal standards ($^{13}C_6$-glucose, D_3-lactate, $^{13}C_3$-pyruvate, D_3-αHB, $^{13}C_2$-βHB, $^{13}C_3$-alanine and $^{13}C_3$-serine) for metabolite extraction. Samples were vortexed and incubated on ice for 30 min. After centrifugation for 20 min at 4 °C at 21,000*g*, 75 μl of sample was transferred to an autosampler glass vial for LC–MS analysis and 10 μl of sample was injected on a Waters XBridge amide column (2.1 × 100 mm, 2.5 μm; part no. 186006091). A pooled QC sample was prepared by mixing an approximately equal volume of each sample and injected every few samples to evaluate the analytical performance. Samples were injected in randomized order to avoid any run order effect. Calibration curves were prepared from 0.5 to 400 μmol/l for pyruvate, αHB, βHB, alanine and serine, 8.32 to 665.6 μmol/l for acetoacetate, and 0.02 to 16 mmol/l for glucose and lactate in surrogate matrix buffer. Human serum albumin (4% w/v) in PBS was used as surrogate matrix buffer. D_3-αHB was used as an internal standard for acetoacetate. The column oven temperature was 27 °C and the autosampler was 4 °C, mobile phase A was 5/95 acetonitrile/water, 20 mM ammonium acetate, pH 9 (adjusted with ammonium hydroxide) and mobile phase B was acetonitrile. The LC gradient conditions at a flow rate of 0.220 ml/min were: 0 min: 85% B, 0.5 min: 85% B, 9 min: 35% B, 11 min: 2% B, 12 min: 2% B, 13.5 min: 85% B, 14.6 min: 85% B, 15 min: 85% B with 0.420 ml/min to 18 min. A Dionex Ultimate 3000 UHPLC system was coupled to a Q-Exactive Plus Orbitrap mass spectrometer (Thermo Fisher Scientific) with a HESI probe operating in polarity switching mode. MS parameters were: a sheath gas flow of 50, an aux gas flow of 10, a sweep gas flow of 2, a spray voltage of 2.50 kV in negative and 3.8 kV in positive, a capillary temperature of 310 °C, a S-lens RF level of −50 and an aux gas heater temperature of 370 °C. Data acquisition was done using Xcalibur software (Thermo Scientific) in the range of 70–1,000 *m/z*, a resolution of 70,000, an automatic gain control (AGC) target of 3×10^6 and a maximum injection time of 80 ms. MS/MS and the retention time of each metabolite were matched against a reference standard to confirm identities. Data analysis was done using Tracefinder 4.1 with 5 ppm mass tolerance, and the quality of integration for each chromatographic peak was reviewed.

For the analysis of liver gluconeogenic intermediates (Extended Data Fig. 7h): using a method adapted from ref. [46], snap-frozen liver tissue (about 40 mg) was ground to a powder using a mortar and pestle on dry ice and extracted using 4/4/2 acetonitrile/methanol/water with 0.1 M formic acid (20 μl of solvent per mg of tissue), vortexed and neutralized with 15% ammonium bicarbonate (8.7 μl for the 100 μl of extraction solvent), sonicated for 1 min and then subjected to two freeze–thaw cycles. Samples were incubated on ice for 20 min and then centrifuged at 21,000*g* for 20 min at 4 °C, at which point 300 μl (30 mg tissue) of supernatant was mixed with 700 μl of water, frozen and lyophilized overnight. Samples were resuspended in 100 μl of 60/40 acetonitrile/water on the day of analysis.

Metabolite separation was performed using a Dionex Ultimate 3000 UHPLC system and ZIC-pHILIC column (150 × 2.1, 5 μm; Merck KGa). Mobile phase A was 20 mM ammonium carbonate in water, pH 9.6 (adjusted with ammonium hydroxide), and mobile phase B was acetonitrile. The column was held at 40 °C, the injection volume was 5 μl and the LC gradient conditions at a flow rate of 0.3 ml/min were: 0 min: 80% B, 0.5 min: 80% B, 20.5 min: 20% B, 21.3 min: 20% B, 21.5 min: 80% B with 7.5 min of equilibration time. MS detection was done with a Q-Exactive Plus Orbitrap mass spectrometer (Thermo Fisher Scientific) with a HESI probe operating in switch polarity mode. MS parameters were: sheath gas flow of 50, an aux gas flow of 12, a sweep gas flow of 2, a spray voltage of 2.80 for negative (3.50 for positive), a capillary temperature of 320 °C, a S-lens RF level of −50 and an aux gas heater temperature of 380 °C. Data acquisition was done using Xcalibur software (Thermo Scientific) and performed in full-scan mode with a range of 70–1,000 *m/z*, a resolution of 70,000, an AGC target of 1×10^6 and a maximum injection time of 80 ms. Data analysis was performed as described above.

For liver NAD(P)(H) analysis (Fig. 2b, Extended Data Fig. 8): ground liver samples were extracted with 4/4/2 acetonitrile/methanol/water with 0.1 M formic acid as described earlier and 5 μl of supernatant was injected on a ZIC-pHILIC column. The column temperature was maintained at 27 °C and the mobile-phase composition was as above. The chromatography gradient was slightly modified: flow rate was

0.15 ml/min, 0 min: 80% B, 0.5 min: 80% B, 20.5 min: 20% B, 21.3 min: 20% B, 21.5 min: 80% B with 7.5 min of column equilibration time.

For liver αHB analysis (Fig. 2c, Extended Data Fig. 3a): liver samples were extracted with D_3-αHB internal standard as described in the gluconeogenic intermediate analysis and analysed on the amide column as described in the media and plasma analysis.

For the sugar analysis by GC–MS (Fig. 5b, mannose): for mannose quantification (Fig. 5b), we used GC–MS rather than LC–MS to separate mannose, glucose, galactose and fructose isomers and quantify them in mouse plasma samples. In brief, 20 µl of plasma sample was extracted with 120 µl of methanol containing 0.83 mM of $^{13}C_3$-glucose and 4.16 µM of $^{13}C_6$-mannose. Samples were vortexed, incubated on ice for 20 min, centrifuged at 21,000*g* at 4 °C for 20 min, and the supernatant was collected. Supernatant (110 µl) was transferred to a glass GC–MS vial and dried down using nitrogen gas. Calibration standards were prepared from 0.5 to 400 µM/l for fructose, mannose, galactose and from 0.02 to 16 mM for glucose in surrogate matrix buffer. We used a two-step derivatization procedure as previously described[47]. First, methoxyamination was performed by adding 50 µl of methoxyamine hydrochloride (20 mg/ml in pyrimidine) to the dried samples and incubating at 30 °C for 90 min. Then, silylation was carried out by adding 80 µl of MSTFA (*N*-methyl-*N*-(trimethylsilyl)trifluoroacetamide) plus 1% TMCS (2,2,2-trifluoro-*N*-methyl-*N*-(trimethylsilyl)-acetamide, chlorotrimethylsilane) and incubating at 70 °C for 60 min. Derivatized samples were cooled down to room temperature before injection. A TriPlus RSH autosampler (Thermo Scientific) was used to inject 1 µl of derivatized sample into a split/splitless (SSL) injector at 250 °C using 1:50 split flow on a TRACE 1310 GC system (Thermo Scientific). Metabolites were separated using a ZB-5MSi GC column (30 m × 0.25 mm × 0.25 µm; Phenomenex). Helium was used as a carrier gas at a flow rate of 1 ml/min. The GC oven program was started at 60 °C and held for 1 min, increased to 320 °C at a rate of 10 °C per min and kept at 320 °C for 5 min. The eluted peaks were transferred through an auxiliary transfer temperature of 300 °C into an electron ionization (EI) source (at 70 eV energy) of the Q-Exactive GC mass spectrometer (Thermo Scientific). High-resolution EI fragmentation spectra were acquired using 60,000 resolution with a mass range of 50–750 *m/z*. The AGC target was 1×10^6 and the maximum injection time was automatic. Data acquisition and analysis were done using Tracefinder 4.1 with 10 ppm mass tolerance and the quality of integration for each chromatographic peak was reviewed.

For the lipidomics measurements (Fig. 5b, Extended Data Fig. 6): lipids were profiled using a Shimadzu Nexera X2 U-HPLC (Shimadzu) coupled to an Exactive Plus Orbitrap mass spectrometer (Thermo Fisher Scientific). In brief, liver tissues (17.0–27.5 mg) were homogenized in four volumes of water (4 µl per mg of tissue) using a bead mill (TissueLyser II, Qiagen). Lipids were extracted from aqueous homogenates (10 µl) using 190 µl of isopropanol containing 1,2-didodecanoyl-*sn*-glycero-3-p hosphocholine (Avanti Polar Lipids). After centrifugation, supernatants were injected directly onto a 100 × 2.1 mm, 1.7 µm ACQUITY BEH C8 column (Waters). The column was eluted isocratically with 80% mobile phase A (95:5:0.1 vol/vol/vol 10 mM ammonium acetate/methanol/formic acid) for 1 min, followed by a linear gradient to 80% mobile phase B (99.9:0.1 vol/vol methanol/formic acid) over 2 min, a linear gradient to 100% mobile phase B over 7 min, then 3 min at 100% mobile phase B. MS analyses were carried out using electrospray ionization in the positive-ion mode using full-scan analysis over 200–1,100 *m/z* at 70,000 resolution and a 3 Hz data acquisition rate. Other MS settings were: sheath gas of 50, in source CID of 5 eV, sweep gas of 5, spray voltage of 3 kV, capillary temperature of 300 °C, S-lens RF of 60, heater temperature of 300 °C, microscans 1, AGC target of 1×10^6 and maximum ion time of 100 ms. Raw data were processed using TraceFinder 3.3 software (Thermo Fisher Scientific) and Progenesis QI (Nonlinear Dynamics). Lipid identities were denoted by total acyl carbon number and total number of double bonds.

Real-time quantitative PCR

RNA was isolated using a Qiagen RNAeasy kit (Qiagen) and reverse transcribed using the SuperScript III First-Strand Synthesis System (Invitrogen). Quantitative PCR was performed using an Applied Biosystems 7500 Fast Real-Time PCR system, using the following primers from Thermo Fisher Scientific: G6Pc (Mm00839363_m), Pepck1 (Mm01247058_m1) and PC (Mm00500992_m1). Hprt (Mm03024075_m1) was used as an internal control. For real-time quantitative PCR experiments, see Extended Data Fig. 7d, g.

Western blotting

Antibody identities and suppliers are available in the Reporting summary.

For western blots in Fig. 2a, Extended Data Figs. 1a, 4c after blotting with primary antibodies, membranes were blotted with HRP-conjugated secondary antibodies and developed with a chemiluminescent substrate (Western Lightning Plus-ECL, PerkinElmer). For remaining western blots, after blotting with primary antibodies, membranes were blotted with IRDye 800CW-conjugated secondary antibodies (Li-Cor) and scanned on a Typhoon bioimager (Amersham). Gels were not reprobed, and each gel band shown is either a different slice from the same gel for bands with different molecular weights, or a separate gel run with the same lysate for those of similar or identical (for example, AKT) molecular weights.

Analysis of NADH/NAD$^+$-sensitive analytes

Metabolites in Fig. 1d were identified as 'NADH/NAD$^+$ sensitive' if the metabolite abundance was significantly different in both the luciferase + ethanol and the *Lb*NOX conditions for a Bonferroni-corrected α of 0.05 using paired, two-tailed Student's *t*-test, and the direction of the changes in metabolite between these two conditions was opposite.

For in vivo experiments, we defined a plasma metabolite as 'NADH/NAD$^+$ sensitive' from the experiment in Fig. 2 if it met all of the following criteria: (i) $P < 0.1$ for two-tailed Student's *t*-test between luciferase and luciferase + EtOH group, (ii) $P < 0.1$ for two-tailed Student's *t*-test between luciferase + EtOH and *Lb*NOX + EtOH group, and (iii) the direction of the changes between these two conditions was opposite.

For the enrichment analysis in Extended Data Fig. 4b, we first determined the redox sensitivity of all metabolites reported in a previous[5] study that we could detect on our metabolomics analysis (156, of which 59 were NADH/NAD$^+$ sensitive). Next, to select independent loci with at least one genome-wide significant ($P < 5 \times 10^{-8}$) metabolite association in ref. [5] (Supplementary Table 2), we clumped all SNPs using the minimum *P* value reported for each SNP, the HapMap CEU reference panel (release 23)[48] and PLINK (v1.9, with options: –clump-p1 5e-8 –clump-p2 0.001 –clump-r2 0.5 –clump-kb 250)[49]. This resulted in 123 independent loci. Next, for each of the 123 loci, we identified all metabolites that are suggestively associated ($P < 1 \times 10^{-5}$) with the lead SNP for that locus, and then performed a one-tailed Fisher's exact test to assess the significance of overlap between these metabolites and the redox-sensitive metabolites identified above.

Table of *GCKR*-associated traits

Supplementary Table 1 was compiled through a combination of manual literature review, the NHGRI-EBI Catalog (https://www.ebi.ac.uk/gwas/), snipa (http://snipa.helmholtz-muenchen.de/snipa3/) and snpedia.com.

Statistics and reproducibility

All data are expressed as mean ± s.e.m. All reported sample sizes (*n*) represent a biologically independent experiment, defined as follows: for all primary hepatocyte data except Peredox experiments (Extended Data Fig. 1c), this was an independent hepatocyte isolation (that is, hepatocytes isolated from a separate mouse) within each experimental condition. For the Peredox experiment, this represents measurements from different individual cells. For in vivo experiments, this represents data from a distinct mouse. For HeLa cell culture experiments, this represents cells seeded on different days within an experimental group. All attempts at replication were successful.

Paired Student's t-tests were used to compare independent hepatocyte isolations divided into experimental groups (for example, *Lb*NOX or luciferase) to account for the significant metabolic variability between isolations from individual mice. All other Student's t-tests were unpaired. For multiple t-tests within an experiment, results were considered significant using a Bonferroni-adjusted $\alpha = 0.05/n$, where n was the number of tests. For ANOVA calculations, a result was considered significant for $P < 0.05$ using a post-hoc Tukey's test.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

All data generated and used in this study are either included in this article (and its Supplementary Information) or are available from the corresponding author on reasonable request.

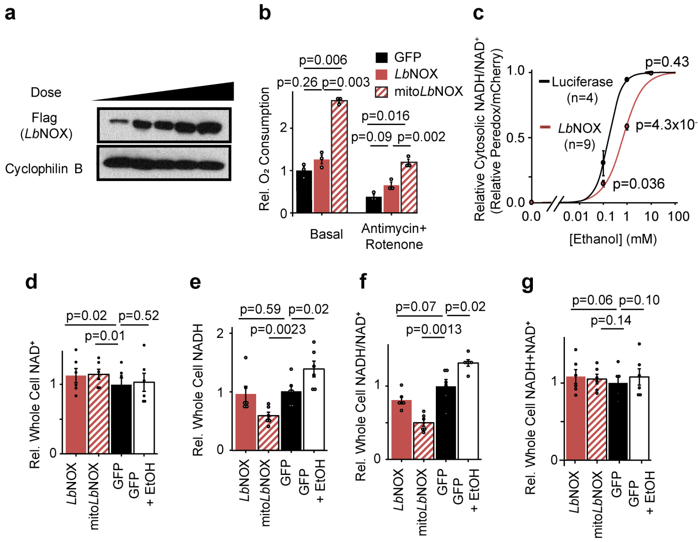
45. Kim, J. K. Hyperinsulinemic-euglycemic clamp to assess insulin sensitivity in vivo. *Methods Mol. Biol*. **560**, 221–238 (2009).
46. Lu, W., Wang, L., Chen, L., Hui, S. & Rabinowitz, J. D. Extraction and quantitation of nicotinamide adenine dinucleotide redox cofactors. *Antioxid. Redox Signal*. **28**, 167–179 (2018).
47. Fiehn, O. Metabolomics by gas chromatography-mass spectrometry: the combination of targeted and untargeted profiling. *Curr. Protoc. Mol. Biol*. **114**, 30.4.1–30.4.32 (2016).
48. International HapMap Consortium et al. The International HapMap Project. *Nature* **426**, 789–796 (2003).
49. Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
50. 1000 Genomes Project Consortium et al. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).

Acknowledgements We thank J. Avruch and members of the Mootha laboratory for valuable discussions and feedback. R.P.G. was supported by the US National Institutes of Health (NIH) grant F32DK111132 and NIH K08DK1158811. R.P.G. and R.M. were both supported by a Pinnacle Research Award from the American Association for the Study of Liver Diseases (AASLD). O.S.S. was supported by NIH F32GM133047. Y.-H.H.H. was supported by NIH T32DK110919. Part of this study was performed at the National Mouse Metabolic Phenotyping Center at University of Massachusetts supported by NIH 5U2C-DK093000 (to J.K.K.) This work was supported in part by a gift from the Marriott Mitochondrial Disorders Collaborative Research Network and grants from the NIH (R35GM122455 and TR01GM099683). V.K.M. is an investigator of the Howard Hughes Medical Institute.

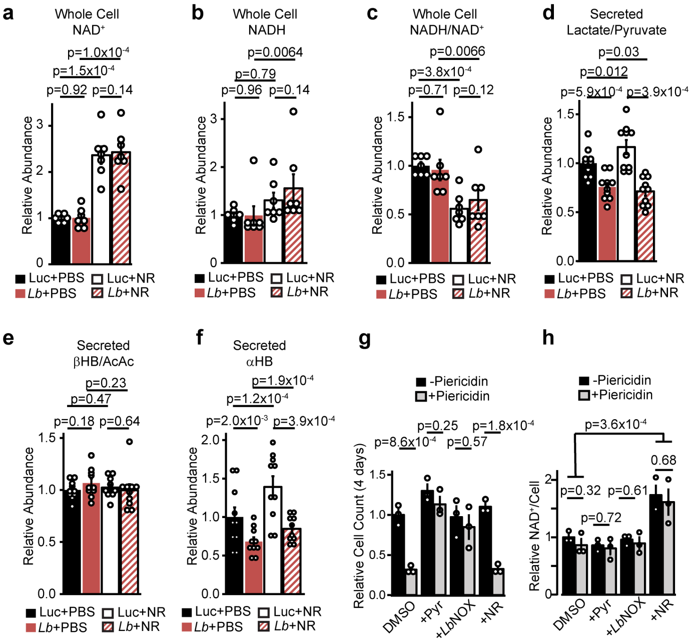
Author contributions R.P.G. and V.K.M. designed the study, which was supervised by V.K.M. R.P.G. and R.M. performed the Peredox experiments, which were supervised by G.Y. R.P.G., A.L.M., A.P. and O.G. performed the in vivo mouse experiments, tissue analyses and hepatocyte experiments. R.P.G., R.S., O.S.S., H.S., C.B.C. and A.D. processed and analysed in vitro and in vivo mouse metabolomics data. H.L.N., S.S. and J.K.K. processed and analysed the hyperinsulinaemic–euglycaemic clamp experiments. Y.-H.H.H. and J.N.H. contributed to the human genetic data analyses. R.P.G. wrote the initial manuscript, which was further edited by V.K.M.

Competing interests V.K.M. is an inventor on the patent PCT/US2016/045015 filed by Massachusetts General Hospital on the use of the *Lb*NOX technology as protein prosthesis for mitochondrial diseases or conditions. V.K.M. and R.P.G. are inventors on a patent provisionally filed by Massachusetts General Hospital on modulating hepatic reductive stress with chemicals.

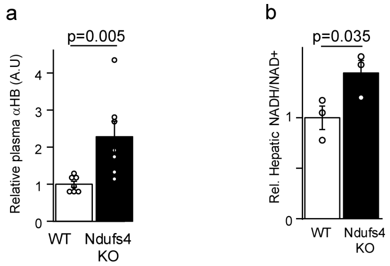
Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2337-2.
Correspondence and requests for materials should be addressed to V.K.M.
Peer review information *Nature* thanks Joshua D. Rabinowitz, Charles M. Brenner and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.
Reprints and permissions information is available at http://www.nature.com/reprints.



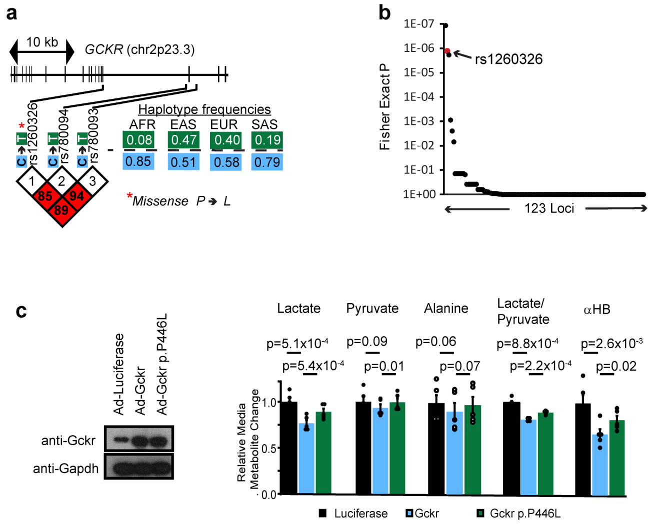
Extended Data Fig. 1 | The effects of *Lb*NOX expression in primary hepatocytes. a, Dose-dependent adenovirus-mediated expression of *Lb*NOX in primary hepatocytes at 24 h. A representative western blot from two independent experiments is shown. **b**, Effect on basal or antimycin + rotenone-insensitive respiration with *Lb*NOX or mito*Lb*NOX. $n = 3$. **c**, Effect of *Lb*NOX on free cytosolic NADH/NAD$^+$ as measured by Peredox with increasing alcohol concentrations. **d**–**g**, Whole cell NAD$^+$ (**d**), NADH (**e**), NADH/NAD$^+$ (**f**) and NADH + NAD$^+$ (**g**). $n = 6$ independent hepatocyte isolations. Nominal P values were determined using paired, two-sided Student's *t*-tests between hepatocyte isolations (**b**, **d**–**g**) or unpaired, two-sided Student's *t*-test for Peredox experiments (**c**). Data are mean ± s.e.m.



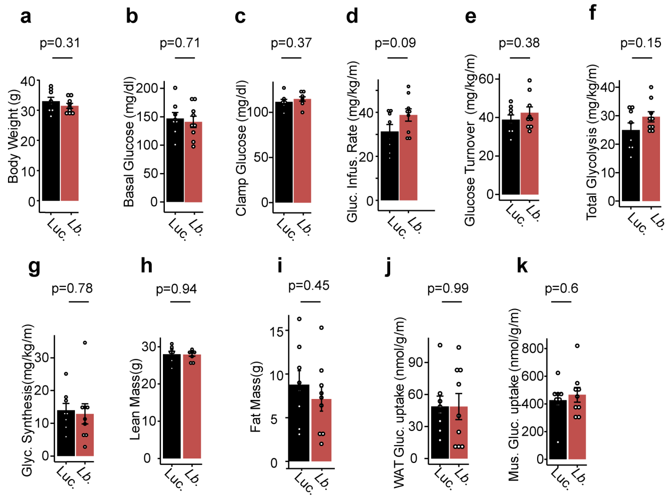
Extended Data Fig. 2 | NR and *Lb*NOX have distinct effects on pyridine dinucleotide pool sizes and redox ratios. a–f, Relative total cellular NAD$^+$ (**a**), NADH (**b**), NADH/NAD$^+$ (**c**), secreted lactate/pyruvate ratio (**d**), secreted βHB/AcAc ratio (**e**) and αHB levels (**f**) in primary hepatocytes with or without NR supplementation and *Lb*NOX. **g**, **h**, Effect of pyruvate (Pyr), *Lb*NOX expression or NR on the inhibition of HeLa cell proliferation (**g**) and total NAD$^+$ levels (**h**) by piericidin 4 days after seeding. Data are mean ± s.e.m. from $n = 7$ (**a**–**c**) or 10 (**d**–**f**) independent hepatocyte isolation, or 3 independent HeLa cell experiments (**g** and **h**). Nominal *P* values were determined using paired (between hepatocyte isolations; **a**–**f**) or unpaired (HeLa cells; **g** and **h**) two-sided Student's *t*-tests.



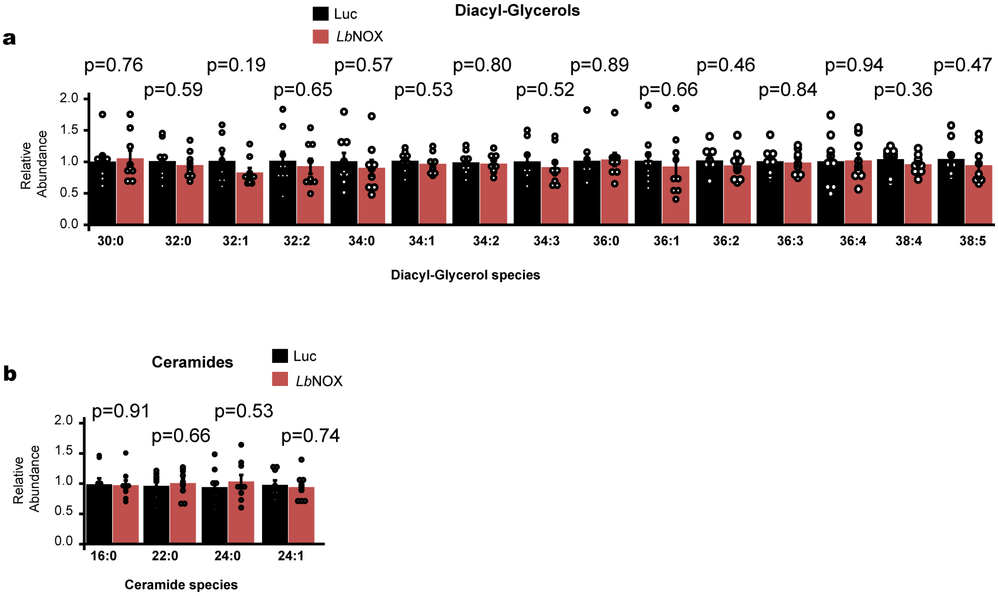
Extended Data Fig. 3 | αHB levels and the hepatic NADH/NAD$^+$ ratio are elevated in the *Ndufs4*-KO mouse model of mitochondrial disease. a, Plasma αHB levels. **b**, Hepatic NADH/NAD$^+$ ratio. $n = 3$ mice in each group for NADH/NAD$^+$ measurements and $n = 7$ mice for αHB measurements. Data are mean ± s.e.m. *P* values were determined using one-sided Student's *t*-test. WT, wild type.



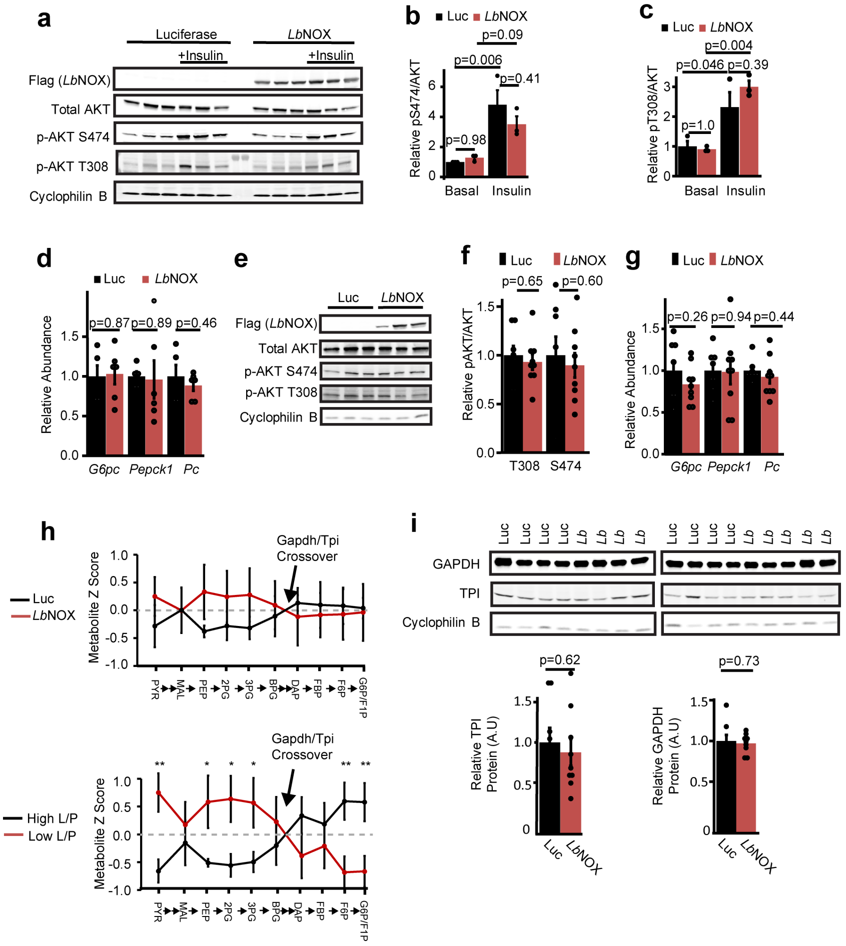
Extended Data Fig. 4 | A common *GCKR* variant increases hepatic cytosolic NADH reductive stress. a, Gene structure, variant location, variant linkage disequilibrium blocks and haplotype frequency for the *GCKR* risk haplotype from the 1000 Genomes Project[50]. AFR, African; EAS, East Asian; EUR, European; SAS, South Asian. **b**, Enrichment of redox-sensitive metabolites in reported metabolite associations by loci from ref. [5]. *P* values were determined from one-tailed Fisher's exact test as described in the Methods. **c**, Effects of overexpression of Gckr or Gckr p.P446L in mouse primary hepatocytes. Data are mean ± s.e.m. Nominal *P* values were determined using paired, two-sided Student's *t*-tests between $n = 5$ independent hepatocyte isolations.



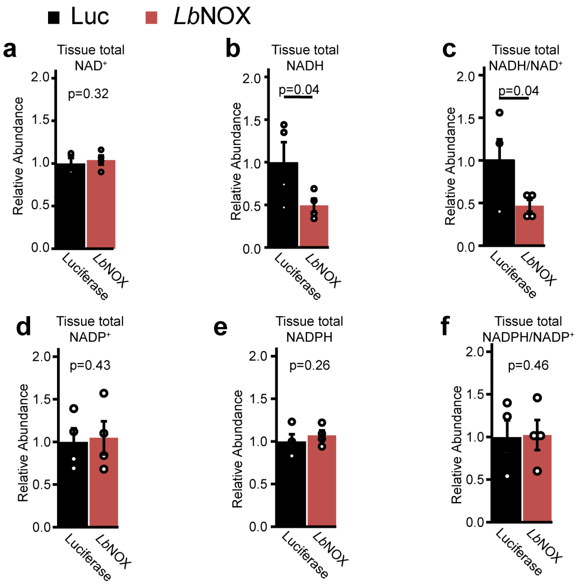
Extended Data Fig. 5 | Effects of *Lb*NOX expression on metabolic parameters during hyperinsulinaemic–euglycaemic clamp. a–**k**, Effects of *Lb*NOX expression in HFD-fed mice using a 2.5 mU min^{-1} kg^{-1} insulin infusion during hyperinsulinaemic–euglycaemic clamp on body weight (**a**), basal glucose levels (**b**), clamp glucose levels (**c**), glucose infusion rate (**d**), whole-body glucose turnover (**e**), whole-body glycolysis (**f**), whole-body glycogen synthesis (**g**), lean mass (**h**), fat mass (**i**), WAT glucose uptake (**j**) and skeletal muscle glucose uptake (**k**). *P* values were determined using two-sided Student's *t*-test. Data are reported as mean ± s.e.m. from $n = 8$ (luciferase) or $n = 9$ (*Lb*NOX) mice.



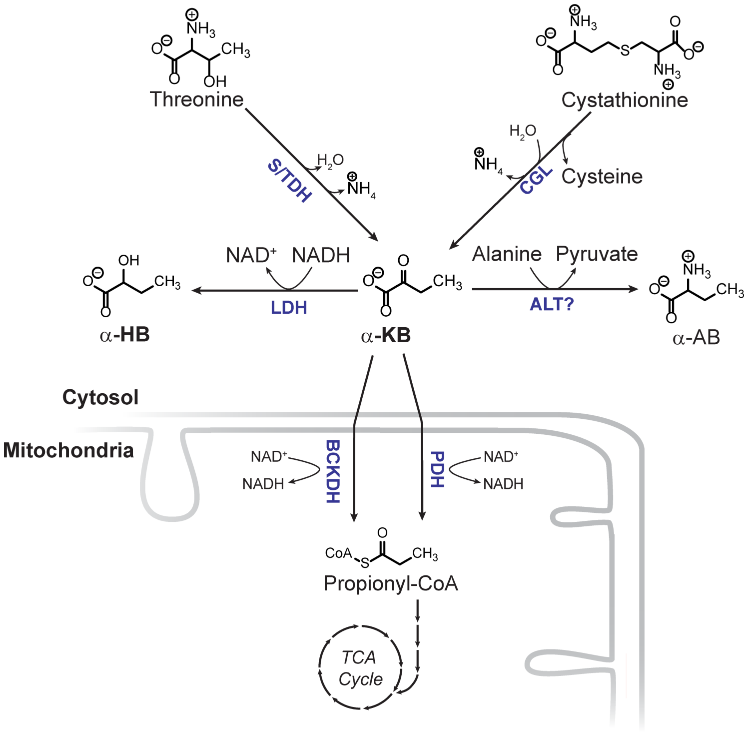
Extended Data Fig. 6 | Hepatic diacyl-glycerols and ceramides in DIO mice with *Lb*NOX expression. a, **b**, Effect of *Lb*NOX expression of hepatic diacyl-glycerol (**a**) and ceramide (**b**) content in DIO mice. Data are reported as mean ± s.e.m. from $n = 9$ mice. *P* values were determined using two-sided Student's *t*-test.



Extended Data Fig. 7 | *Lb*NOX improves hepatic insulin resistance in vivo independent of hepatic insulin signalling. a–c, Western blots of liver lysate from DIO mice 15 min after an intraperitoneal injection of saline or 2 U/kg insulin (**a**) with relative pS474 AKT/total AKT (**b**) and relative pT308 AKT/total AKT (**c**). $n = 3$ from representative western blots from 2 independent experiments. **d**, Transcriptional FOXO1 targets *G6pc*, *Pepck1* and *Pc* in DIO mice with *Lb*NOX or luciferase. $n = 6$. **e**, Western blots of liver lysates at the end of hyperinsulinaemic–eugylcaemic clamps. $n = 3$ representative of $n = 8$ (luciferase) and 9 (*Lb*NOX). **f**, **g**, Relative pS474 AKT/total AKT and pT308 AKT/total AKT (**f**) and transcriptional FOXO1 targets *G6pc*, *Pepck1* and *Pc* (**g**). $n = 8$ (luciferase) and 9 (*Lb*NOX). **h**, Crossover analysis of relative abundance of gluconeogenic intermediates at the end of hyperinsulinaemic–euglycaemic clamps. Top, *Lb*NOX versus Luc mice are compared. Bottom, samples are divided by high or low liver lactate/pyruvate (L/P) ratios and compared. $^*P < 0.05$, $^{**}P < 0.01$, using two-sided Student's *t*-test. BPG, 1,3-bisphosphoglycerate; DAP, dihydroxyacetone phosphate; FBP, fructose 1,6-bisphosphate; F1P, fructose 1-phosphate; F6P, fructose 6-phosphate; G6P, glucose 6-phosphate; MAL, malate; PEP, phosphoenolpyruvate; 2PG, 2-phosphoglycerate; 3PG, 3-phosphoglycerate; PYR, pyruvate. **i**, Western blots and relative protein levels of GAPDH and triosephosphate isomerase (TPI) at the end of the insulin clamp. $n = 8$ (luciferase) and 9 (*Lb*NOX). Data are reported as mean ± s.e.m.



Extended Data Fig. 8 | NAD(P)(H) levels in *Lb*NOX versus luciferase livers at the end of hyperinsulinaemic–euglycaemic clamps. a–**f**, The relative abundance of total NAD$^+$ (**a**), total NADH (**b**), total NADH/NAD$^+$ (**c**), total NADP$^+$ (**d**), total NADPH (**e**) and total NADPH/NADP$^+$ (**f**). $n = 4$ mice for each group. P values were determined by one-sided Student's t-test. Data are reported as mean ± s.e.m.



Extended Data Fig. 9 | Metabolic origins and fate of αHB. α-AB, α-aminobutyrate; BCKDH, branched-chain α-keto acid dehydrogenase complex; CGL, cystathionine γ-lyase; PDH, pyruvate dehydrogenase; S/TDH, serine/threonine dehydratase; TCA, tricarboxylic acid.

nature research

Corresponding author(s): Vamsi Mootha
Manuscript: 2019-02-02819A

Last updated by author(s): Nov 10, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Xcalibur (v. 4.1.31.9, Thermofisher), Amsheram Typhoon (1.0.0.7, GE Healthcare), iQ (v 2.9.1, Andor)
Data analysis	Microsoft Excel for Office 365 MSO, R (v 3.6.2), ImageJ (v 1.51, NIH), Tracefinder (v 4.1, Thermofisher) and Progenesis (2.3.6275.47961, nonlinear dynamics).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Statistical tests were not used to predefine sample size. For insulin clamp and related experiments, sample sizes of 8-10 were chosen to demonstrate moderate differences in commonly measured metabolic parameters, and because these are typical sample sizes reported for such experiments in the literature. For the in vivo LbNOX/EtOH experiments and subsequent metabolomics experiments, sample of n=8 were targeted as these numbers were sufficient in hepatocyte experiments to demonstrate differences in secreted aHB levels.
Data exclusions	For in vivo mouse experiments involving adenovirus-mediated expression of LbNOX or control, mice were not included in experiments which failed due to technical reasons. A single sample was excluded from figure S3H and from Figure 6B (urate measurements) as they were outliers as determined by Grubbs statistical tests. These criteria were not predefined before performing the experiment.
Replication	All attempts at replication were successful. All data points represent either experiments performed on an individual mouse (for in vivo data), hepatocytes isolated from an individual mouse (for data points in each experimental group in primary hepatocyte data), or an experiment performed on a separate day (for data in each experimental group in cell line experiments).
Randomization	For in vivo experiments involving LbNOX or luciferase injections, cagemates were were randomly assigned to a particular experimental condition. Mass spec runs involved randomization of sample order.
Blinding	Blinding was not performed for experiments described in this paper.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	1. GAPDH (Cell Signaling Technology, Catalog #2118, Clone 14C10, Lot 1, 1:1000 Dilution) 2. Flag (Cell Signaling Technology, Catalog #2368, Polyclonal, Lot 12, 1:1000 Dilution) 3. Pan AKT (Cell Signaling Technology, Catalog #4691, Clone C67E7, Lot 20, 1:1000 Dilution) 4. Phospho-AKT2 Ser474 (Cell Signaling Technology, Catalog #8599, Clone D3H2, Lot 2, 1:1000 Dilution) 5. Phospho-AKT Thr308 (Cell Signaling Technology, Catalog #13038, Clone D25E6, Lot 5, 1:1000 Dilution) 6. Cyclophilin B (Abcam, Catalog #ab178397, Clone EPR12703(b), LotGR317882-10, 1:500 Dilution) 7. GCKR (Cell Signaling Technology, Catalog #14328, Clone D1W9P, Lot 1, 1:1000 Dilution) 8. TPI aka TIM (Santa Cruz Biotechnology, Catalog #sc-166785, Clone H-11, Lot#B2117, Dilution 1:200)
Validation	GCKR and Flag antibodies were validated in this paper through detection of adenoviral overexpression of targets of the appropriate molecular weight. pan-AKT was validated by the supplier through detection of recombinant AKT1, AKT2, and AKT3. phospho-AKT2 ser474 was validated by the supplier in genetic knockout of endogenous AKT2 in mouse embryonic fibroblasts. Phospho-AKT Thr308 was not validated by the supplier in a genetic knockout, but has been shown to detect AKT in the context of hPDGF exposure in human cell lines. Cyclophilin B was validated by the supplier in a genetic knockout of endogenous protein in a human cell line. GAPDH and TPI were not validated by the supplier or in this paper.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Hela cells were obtained from ATCC, and then modified as previously described (see Titov et al. Science 2016 DOI: 10.1126/science.aad401). 293A for adenovirus amplification were also obtained from ATCC.
Authentication	The cell lines used in this paper from ATCC were not specifically re-authenticated.
Mycoplasma contamination	All cell lines were tested for mycoplasma contamination monthly which was negative.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified lines were used in this paper.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	All animals used were male C57BL6/J mice purchased from Jackson labs, aged 12-16 weeks.
Wild animals	This study did not involve wild animals.
Field-collected samples	This study did not involve field-collected samples.
Ethics oversight	All animal protocols were approved by the MGH IACUC

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Senolytic CAR T cells reverse senescence-associated pathologies

https://doi.org/10.1038/s41586-020-2403-9

Received: 24 September 2019

Accepted: 6 May 2020

Corina Amor[1,2,12], Judith Feucht[3,4,12], Josef Leibold[2,12], Yu-Jui Ho[2], Changyu Zhu[2], Direna Alonso-Curbelo[2], Jorge Mansilla-Soto[3,4], Jacob A. Boyer[1,5], Xiang Li[2,6], Theodoros Giavridis[3,4], Amanda Kulick[5], Shauna Houlihan[2], Ellinor Peerschke[7], Scott L. Friedman[8], Vladimir Ponomarev[9], Alessandra Piersigilli[10], Michel Sadelain[3,4]✉ & Scott W. Lowe[2,11]✉

Cellular senescence is characterized by stable cell-cycle arrest and a secretory program that modulates the tissue microenvironment[1,2]. Physiologically, senescence serves as a tumour-suppressive mechanism that prevents the expansion of premalignant cells[3,4] and has a beneficial role in wound-healing responses[5,6]. Pathologically, the aberrant accumulation of senescent cells generates an inflammatory milieu that leads to chronic tissue damage and contributes to diseases such as liver and lung fibrosis, atherosclerosis, diabetes and osteoarthritis[1,7]. Accordingly, eliminating senescent cells from damaged tissues in mice ameliorates the symptoms of these pathologies and even promotes longevity[1,2,8–10]. Here we test the therapeutic concept that chimeric antigen receptor (CAR) T cells that target senescent cells can be effective senolytic agents. We identify the urokinase-type plasminogen activator receptor (uPAR)[11] as a cell-surface protein that is broadly induced during senescence and show that uPAR-specific CAR T cells efficiently ablate senescent cells in vitro and in vivo. CAR T cells that target uPAR extend the survival of mice with lung adenocarcinoma that are treated with a senescence-inducing combination of drugs, and restore tissue homeostasis in mice in which liver fibrosis is induced chemically or by diet. These results establish the therapeutic potential of senolytic CAR T cells for senescence-associated diseases.

Given the contribution of senescence to tissue damage, there is growing interest in the development of 'senolytic' agents that selectively eliminate senescent cells[12]. Several small molecules exhibit senolytic activity, but most lack potency and produce substantial side effects[1,12,13]. An alternative approach could involve CAR T cells directed against senescence-specific surface antigens. CARs are synthetic receptors that redirect T cell specificity, effector potential and other functions[14]. CAR T cells that target CD19 have shown notable efficacy in patients with refractory B cell malignancies[15], and other cell-surface antigens show promise as targets for CAR therapy in different contexts[16–18]. Here we investigate whether CAR T cells could serve as senolytic agents.

Upregulation of uPAR during senescence

To identify cell-surface proteins that are broadly and specifically upregulated in senescent cells, we compared RNA-sequencing (RNA-seq) datasets derived from three independent and robust models of senescence: 1) therapy-induced senescence in mouse lung adenocarcinoma *Kras*^{G12D};*p53*$^{-/-}$ (KP) cells (*p53* is also known as *Trp53*) that are triggered to senesce by a combination of MEK inhibition and CDK4 and CDK6 (CDK4/6) inhibition[19]; 2) oncogene-induced senescence in mouse hepatocytes, mediated by the in vivo delivery of *Nras*^{G12V} through hydrodynamic tail vein injection (HTVI)[4]; and 3) culture-induced senescence in mouse hepatic stellate cells (HSCs) (Extended Data Fig. 1a). We focused on transcripts that encode molecules that are located in the plasma membrane (as determined by UniProtKB) and that were upregulated in all datasets (Extended Data Fig. 1b, c). Eight transcripts were identified, which encode proteins related to extracellular matrix remodelling or the coagulation cascade (Extended Data Fig. 1d).

Given that ideal antigens for the engagement of CAR T cells should be highly expressed on target cells but not in vital tissues, we ranked each transcript according to its magnitude of upregulation ($\log_2$(expression in senescent cells/expression in non-senescent cells)), and then excluded those that were highly expressed in vital tissues (as determined by the Human Protein Atlas and the Human Proteome Map)[20]. This process identified *PLAUR*, which encodes the

---

[1]Louis V. Gerstner Jr Graduate School of Biomedical Sciences, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [2]Department of Cancer Biology and Genetics, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [3]Center for Cell Engineering, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [4]Immunology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [5]Molecular Pharmacology and Chemistry Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [6]Weill Cornell Graduate School of Medical Sciences, New York, NY, USA. [7]Department of Laboratory Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [8]Division of Liver Diseases, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [9]Department of Radiology, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [10]Laboratory of Comparative Pathology, Rockefeller University, Weill Cornell Medicine and Memorial Sloan Kettering Cancer Center, New York, NY, USA. [11]Howard Hughes Medical Institute, Chevy Chase, MD, USA. [12]These authors contributed equally: Corina Amor, Judith Feucht, Josef Leibold. ✉e-mail: m-sadelain@ski.mskcc.org; lowes@mskcc.org

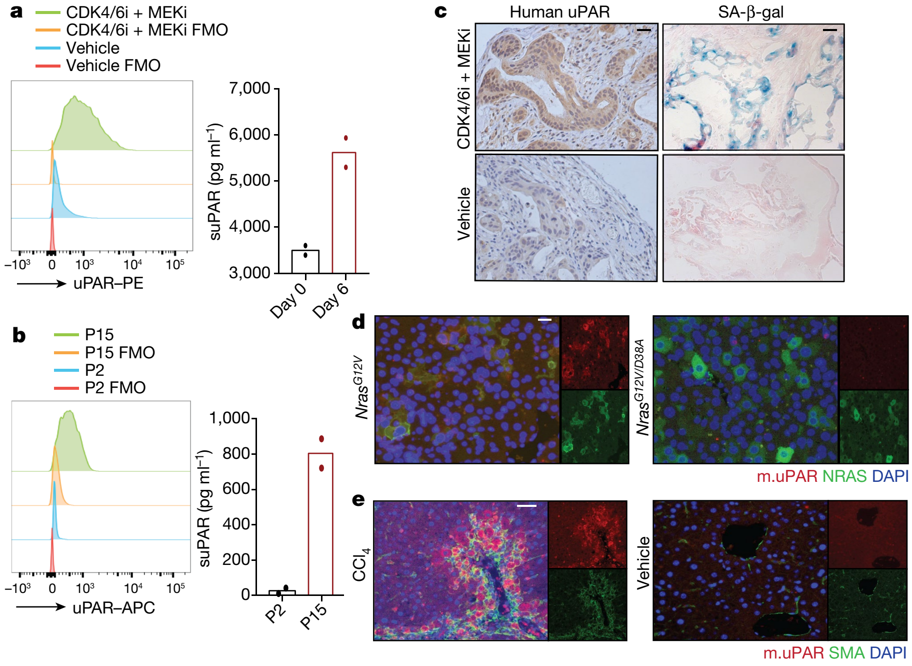


Fig. 1 | uPAR is a cell-surface and secreted biomarker of senescence. a, Left, flow cytometry analysis of uPAR expression on mouse KP lung adenocarcinoma cells after induction of senescence by treatment with MEK and CDK4/6 inhibitors (CDK4/6i + MEKi) as compared to controls. FMO, fluorescence minus one control. Representative results of $n = 3$ independent experiments. Right, levels of suPAR as determined by enzyme-linked immunosorbent assay (ELISA) in the supernatant of senescent or proliferating KP cells. Representative results of $n = 2$ independent experiments. **b**, Left, flow cytometry analysis comparing uPAR expression on primary human melanocytes after induction of senescence by continuous passage (P) with proliferating controls. Representative results of $n = 2$ independent experiments. Right, levels of suPAR in the supernatant of senescent (passage 15; P15) or proliferating (passage 2; P2) primary human melanocytes. Representative results of $n = 2$ independent experiments. **c**, Immunohistochemical staining of human uPAR and SA-β-gal in a patient-derived xenograft from human lung adenocarcinoma orthotopically injected into NSG mice and treated with vehicle or combined MEK and CDK4/6 inhibitors. Representative of $n = 2$ independent experiments ($n = 3$ mice per group). Scale bars, 50 µm. **d**, Co-immunofluorescence staining of uPAR (red) and NRAS (green) in the livers of mice six days after transfection by HTVI with a plasmid encoding $Nras^{G12V}$ or $Nras^{G12V/D38A}$. Representative results of $n = 3$ independent experiments ($n = 5$ mice per group). Scale bar, 50 µm. **e**, Co-immunofluorescence staining of uPAR (red) and smooth muscle actin (SMA; green) in the livers of mice six weeks after intraperitoneal treatment twice weekly with CCl_4 ($n = 7$ mice) or vehicle ($n = 4$ mice). Representative results of $n = 3$ independent experiments. Scale bar, 50 µm.

urokinase-type plasminogen activator receptor (uPAR), as a suitable candidate (Extended Data Fig. 1e). Accordingly, *PLAUR* was also upregulated in public datasets of senescent human cells[19,21] and immunohistochemistry confirmed that uPAR protein was absent in many vital organs (Extended Data Fig. 1g, f). Consistent with previous reports, low uPAR expression was detected in the bronchial epithelium. Other cell types that express uPAR include subsets of monocytes, macrophages and neutrophils[11,22].

uPAR is the receptor for urokinase-type plasminogen activator, which promotes the degradation of the extracellular matrix during fibrinolysis, wound healing or tumorigenesis[11]. uPAR also functions as a signalling receptor that promotes the motility, invasion and survival of tumour cells[11]. Nonetheless, mice that lack uPAR are viable and fertile[23]. A portion of uPAR is proteolytically cleaved upon ligand binding, which generates soluble uPAR (suPAR). Notably, suPAR is secreted by senescent cells as part of the senescence-associated secretory phenotype (SASP)[24] and serves as a serum biomarker for kidney disease and diabetes[25]—two chronic pathologies that are linked to senescence[25].

We next confirmed that uPAR expression was induced on the surface of senescent cells in vitro and in vivo. First, we evaluated therapy-induced senescence in mouse KP lung cancer cells that were treated with combined MEK and CDK4/6 inhibition, and replication-induced senescence in human primary melanocytes (Fig. 1a, b, Extended Data Fig. 2a, b). In both models, cell-surface expression of uPAR and supernatant suPAR levels were markedly increased after the induction of senescence (Fig. 1a, b). Second, we examined a patient-derived xenograft model of non-small-cell lung cancer in which mice were treated with combined MEK and CDK4/6 inhibitors[19] (Fig. 1c) and two different models of oncogene-induced senescence triggered either by the overexpression of $Nras^{G12V}$ in mouse hepatocytes transfected by HTVI (Fig. 1d, Extended Data Fig. 2c–e) or by the endogenous expression of $Kras^{G12D}$ in a mouse model of senescent pancreatic intraepithelial neoplasia (Extended Data Fig. 2f–i). Finally, we included a mouse model of carbon tetrachloride (CCl_4)-induced liver fibrosis, in which senescent HSCs contribute to the pathophysiology[6] (Fig. 1e, Extended Data Fig. 2j–m). In each system, the senescence-inducing treatment led to an increase in the number of uPAR-positive cells and an increase in serum suPAR levels. Notably, uPAR-positive cells did not express the proliferation marker Ki-67, but co-expressed interleukin 6 (IL-6)—an established component of the SASP[1,2].

We next confirmed that uPAR is highly expressed in tissues from patients with senescence-associated disorders. High levels of uPAR expression were observed in specimens of liver fibrosis of different aetiologies. uPAR-positive cells showed the same histological presentation as cells that expressed senescence-associated β-galactosidase (SA-β-gal), and co-expressed the senescence-associated markers p16 and IL-6 (Extended Data Fig. 3a, b). uPAR was also highly expressed in atherosclerotic plaques from human carotid endarterectomy specimens and in pancreatic intraepithelial neoplasia lesions from patients with pancreatic cancer (Extended Data Fig. 3c, d). In addition, increased levels of uPAR and/or suPAR have been noted in patients with other diseases that are associated with senescence, including osteoarthritis, diabetes and idiopathic pulmonary fibrosis[26–28]. Collectively, these results show that uPAR is a candidate target for senolytic CAR T cells.

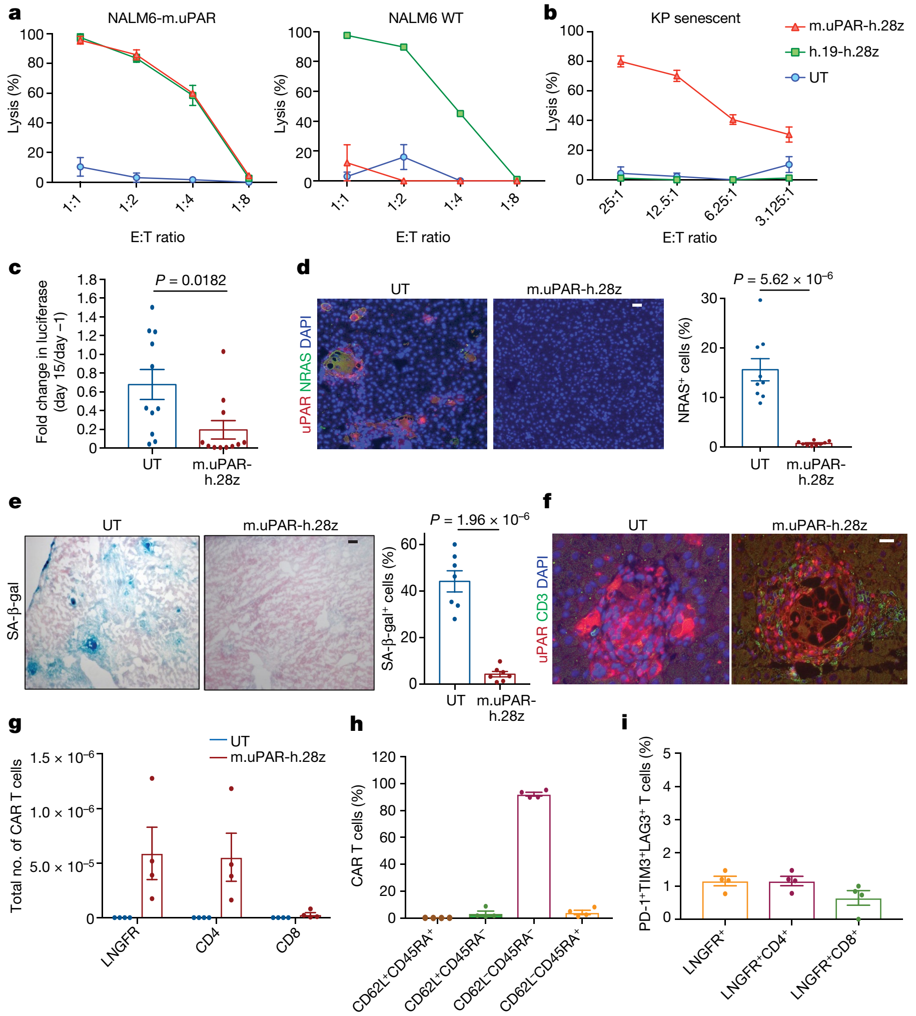


Fig. 2 | uPAR CAR T cells are bona fide senolytic agents. a, Cytotoxic T cell activity as determined by an 18-h bioluminescence assay using luciferase-expressing wild-type (WT) NALM6 cells or NALM6 cells that overexpress mouse uPAR (NALM6-m.uPAR) as targets. E:T ratio, effector-to-target ratio; UT, untransduced T cells. Data are representative of $n = 3$ independent experiments, each performed in triplicate. **b**, Cytotoxic T cell activity as determined by a 4-h bioluminescence assay using KP cells as targets in which senescence was induced by MEK and CDK4/6 inhibition. Data are representative of $n = 2$ independent experiments, each performed in triplicate. **c**–**i**, NSG mice were injected with a plasmid encoding Nras^{G12V}-GFP-luciferase and treated with 0.5×10^6 m.uPAR-h.28z CAR T cells or untransduced T cells 10 days after injection. Mice were euthanized 15 days later and livers were analysed. **c**, Fold change in luciferase signal in mice (calculated as the average radiance on day 15 divided by the average radiance on day −1) ($n = 11$ mice per group). **d**, Co-immunofluorescence staining of uPAR (red) and NRAS (green) and quantification of NRAS-positive cells ($n = 9$ mice per group). Scale bar, 50 μm. **e**, Representative staining and quantification of SA-β-gal-positive cells ($n = 7$ mice per group). Scale bar, 50 μm. **f**, Co-immunofluorescence staining of uPAR (red) and human CD3 (green), showing T cell infiltration ($n = 5$ mice per group). Scale bar, 50 μm. **g**–**i**, Number of liver-infiltrating CAR T cells (**g**), expression of CD62L and CD45RA (**h**) and percentage of PD-1$^+$TIM3$^+$LAG3$^+$ CAR T cells (**i**) among m.uPAR-h.28z CAR T cells as determined by flow cytometry ($n = 4$ mice per group). Representative results of $n = 2$ independent experiments (**c**–**f**). Data are mean ± s.e.m.; two-tailed unpaired Student's *t*-test (**c**–**e**).

Senolytic activity of uPAR CAR T cells

We constructed a uPAR-specific CAR comprising an anti-mouse uPAR (m.uPAR) single-chain variable fragment linked to human CD28 costimulatory and CD3ζ (h.28z) signalling domains (m.uPAR-h.28z), transduced human T cells and performed cytotoxicity assays using target cells that express a mouse uPAR cDNA (Extended Data Fig. 4a–d). To enable comparisons to well-characterized CAR T cells directed against CD19[29], mouse uPAR was introduced into the human CD19$^+$ pre-B acute lymphoblastic leukaemia (B-ALL) cell line NALM6 (Extended Data Fig. 4c). m.uPAR-h.28z CAR T cells showed no cytotoxicity towards uPAR-negative NALM6 cells, but comparable activity to CD19-specific CAR T cells incorporating human CD28 and CD3ζ signalling elements (h.19-h.28z) when targeting uPAR-expressing NALM6 cells (Fig. 2a, Extended Data Fig. 4d). m.uPAR-h.28z—but not h.19-h.28z—CAR T cells efficiently eliminated senescent KP cells that express endogenous uPAR, and this was accompanied by antigen-specific secretion of granzyme B and interferon γ (IFNγ) (Fig. 2b, Extended Data Fig. 4e). Hence, m.uPAR-h.28z CAR T cells can selectively and efficiently target senescent cells.

To study whether m.uPAR-h.28z CAR T cells could function as a bona fide senolytic agent in vivo, we took advantage of the well-characterized model of oncogene-induced senescence triggered by the hepatic overexpression of Nras^{G12V}-luciferase used above[4]. Although these senescent cells normally undergo SASP-mediated immune clearance[4], they are retained in the livers of immunodeficient NOD scid gamma (NSG) mice[4]. Successful transfection of *Nras*^{G12V} into hepatocytes of NSG mice was confirmed by bioluminescence imaging, and was followed by administration of 0.5×10^6 m.uPAR-h.28z CAR T cells or untransduced T cells as controls (Extended Data Fig. 4f).

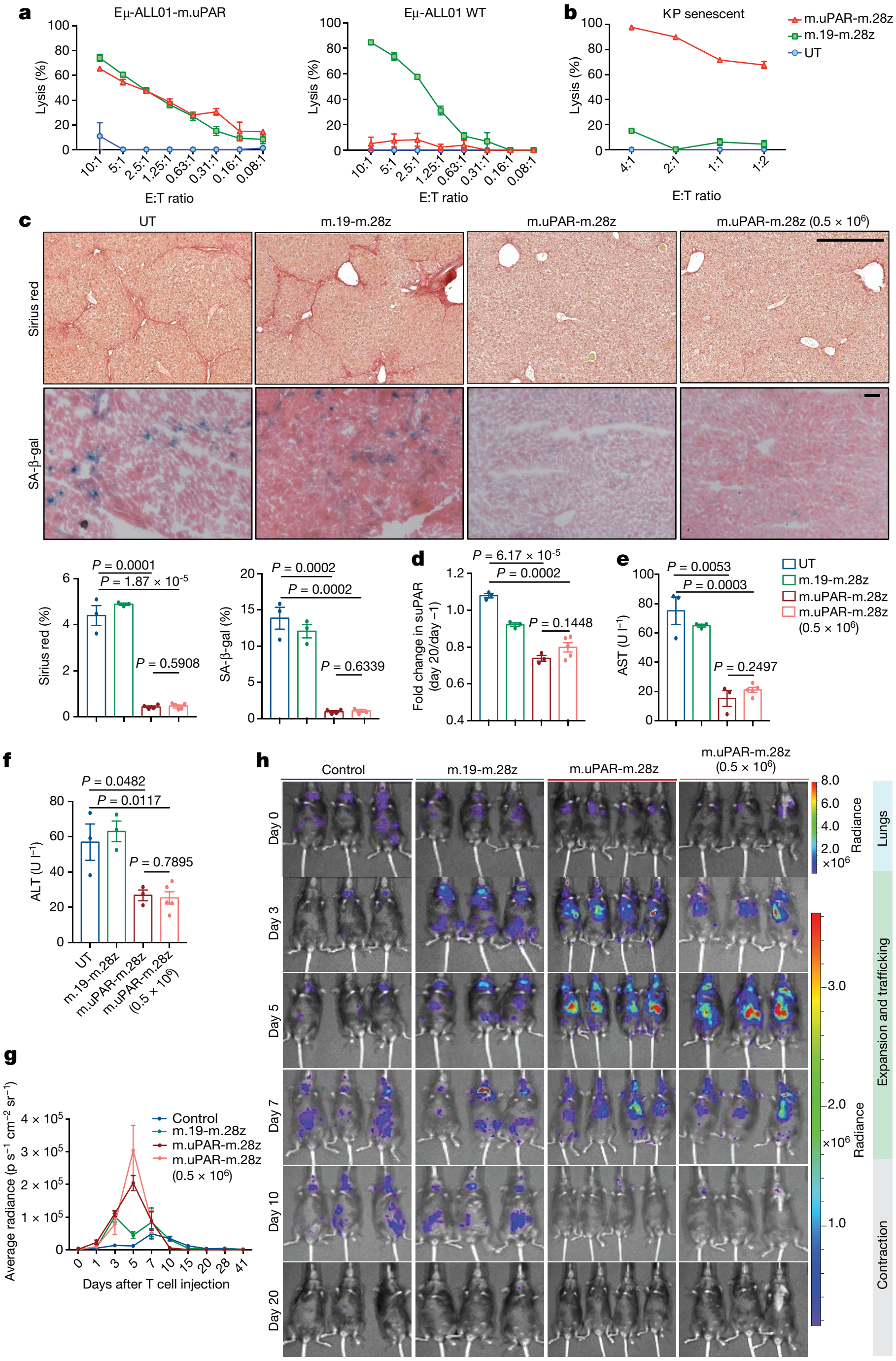


Fig. 3 | Senolytic CAR T cells show therapeutic efficacy in CCl_4-induced liver fibrosis. a, Cytotoxicity of mouse CAR T cells as determined by an 18-h bioluminescence assay using luciferase-expressing wild-type Eμ-ALL01 cells (WT) or Eμ-ALL01 cells that overexpress mouse uPAR (Eμ-ALL01-m.uPAR) as targets. Data are representative of $n = 3$ independent experiments, each performed in triplicate. **b**, Cytotoxic T cell activity as determined by an 18-h bioluminescence assay using KP cells as targets in which senescence was induced by MEK and CDK4/6 inhibition. Data are representative of $n = 2$ independent experiments, each performed in triplicate. **c**–**f**, Mice with CCl_4-induced liver fibrosis were treated with 0.5×10^6 or 1×10^6 m.uPAR-m.28z CAR T cells, 1×10^6 m.19-m.28z CAR T cells or untransduced T cells and euthanized 20 days later. Livers were used for further analyses. **c**, Representative levels of fibrosis as evaluated by Sirius red staining and SA-β-gal expression (top) and respective quantifications (bottom) (UT and m.19-m.28z, $n = 3$; m.uPAR-m.28z, $n = 4$; m.uPAR-m.28z at 0.5×10^6, $n = 5$ mice). Scale bars, 500 μm (top); 50 μm (bottom). **d**, Fold change difference in the serum levels of suPAR 20 days after (day 20) compared to 1 day before (day −1) infusion of T cells. **e**, **f**, Levels of serum AST (**e**) and ALT (**f**) 20 days after infusion of T cells (UT, m.19-m.28z and m.uPAR-m.28z, $n = 3$; m.uPAR-m.28z at 0.5×10^6, $n = 5$ mice (**d**–**f**)). **g**, **h**, Mice with CCl_4-induced liver fibrosis were injected with 0.5×10^6 or 1×10^6 m.uPAR-m.28z CAR T cells, 1×10^6 m.19-m.28z CAR T cells or control T cells that were transduced to express click beetle red luciferase. **g**, Luciferase signal (average radiance) of treated mice after administration of T cells, reflecting the expansion of T cells (control T cells and m.19-m.28z, $n = 3$; m.uPAR-m.28z, $n = 4$; m.uPAR-m.28z at 0.5×10^6, $n = 3$ mice). **h**, Representative bioluminescence images of mice at different time points after injection. T cells were initially detected in the lungs in all treated mice; m.uPAR-m.28z CAR T cells showed trafficking to the liver area followed by a short period of expansion and a rapid contraction. The signal in control mice at day 10 indicates abdominal peritonitis induced by CCl_4 injections, as confirmed by pathology. For the colour scales on the right (measured in p s^{-1} cm^{-2} sr^{-1}), the minimum value and maximum values are 1.43×10^4 and 8.00×10^6, respectively (top) and 1.50×10^5 and 3.63×10^6, respectively (bottom). Results of $n = 1$ independent experiment (**c**–**h**). Data are mean ± s.e.m.; two-tailed unpaired Student's *t*-test (**c**–**f**).

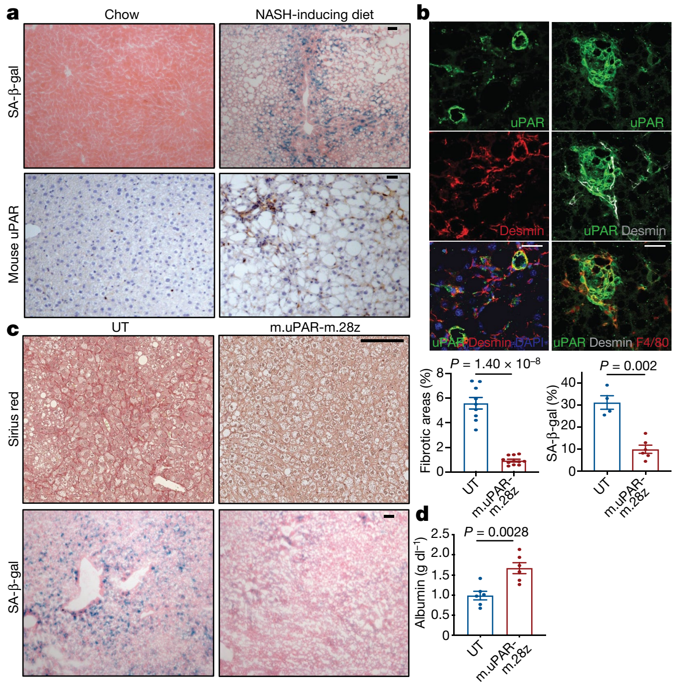


Fig. 4 | Senolytic CAR T cells are therapeutic in NASH-induced liver fibrosis. **a**, **b**, Representative staining in the livers of mice that were treated with a chow or a NASH-inducing diet for 3–4 months. **a**, Immunohistochemical staining of mouse uPAR and SA-β-gal. Scale bars, 50 μm. **b**, Co-immunofluorescence staining of mouse uPAR (green), desmin (red, left; or grey, right) and F4/80 (red) in livers of mice treated with NASH-inducing diet for 3–4 months. Representative results of $n = 2$ independent experiments ($n = 3$ mice per group). Scale bars, 30 μm. **c**, **d**, Mice that were treated with a NASH-inducing diet for 3 months were injected with 0.5×10^6 m.uPAR-m.28z CAR T cells or untransduced T cells. Liver and serum analyses were performed 20 days later. **c**, Representative images of Sirius red staining and SA-β-gal expression (left) and quantifications (right) (Sirius red: UT, $n = 9$; m.uPAR-m.28z, $n = 11$; SA-β-gal: UT, $n = 4$; m.uPAR-m.28z, $n = 6$ mice). Scale bars, 200 μm (top); 50 μm (bottom). **d**, Serum albumin levels (UT and m.uPAR-m.28z, $n = 6$ mice). Results of $n = 2$ independent experiments (**c**, **d**). Data are mean ± s.e.m.; two-tailed unpaired Student's *t*-test (**c**, **d**).

Treatment with m.uPAR-h.28z CAR T cells led to a profound decrease in the bioluminescence signal within 10 days (Fig. 2c), suggesting effective clearance of senescent hepatocytes. Histological analyses confirmed that livers from mice that were treated with m.uPAR-h.28z CAR T cells had significantly reduced numbers of NRAS-positive cells ($P < 0.01$) and SA-β-gal-positive cells ($P < 0.001$) compared to livers from control mice (Fig. 2d, e). Furthermore, m.uPAR-h.28z CAR T cells (but not untransduced T cells) accumulated around senescent hepatocytes within 7 days of infusion (Fig. 2f) and displayed an effector memory phenotype ($CD62L^-CD45RA^-$) with little evidence of T cell exhaustion (less than 2% $PD\text{-}1^+TIM3^+LAG3^+$ CAR T cells) 15 days after their administration (Fig. 2g–i, Extended Data Fig. 11). Therefore, uPAR-28z CAR T cells can eliminate senescent cells in vivo.

Efficacy of senolytic CAR T cells

To evaluate the senolytic capacity of uPAR CAR T cells in immunocompetent settings, we transduced T cells from C57BL/6 mice with a fully mouse CAR (m.uPAR-m.28z). We confirmed CAR expression and excluded uPAR expression on transduced T cells, and showed that they exhibited a similar cytolytic profile to m.uPAR-h.28z CAR T cells when targeting the mouse $CD19^+$ B-ALL cell line Eμ-ALL01 genetically modified to express exogenous uPAR or when targeting senescent KP cells (Fig. 3a, b, Extended Data Fig. 4g, h).

Previous studies suggest that the combination of a senescence-inducing cancer therapy and a senolytic agent can improve treatment outcome in mouse models[30]. We thus treated mice that had orthotopic KP lung adenocarcinomas with combined MEK and CDK4/6 inhibitors[19], and then administered uPAR- or CD19-specific CAR T cells or untransduced T cells (Extended Data Fig. 5a). Treatment with uPAR-targeted CAR T cells significantly prolonged survival without eliciting signs of toxicity (Extended Data Fig. 5b–d). Lungs that were collected from mice treated with uPAR-specific CAR T cells showed a substantial decrease in senescent tumour cells, accompanied by enhanced infiltration of adoptively transferred $CD4^+$ and $CD8^+$ T cells that expressed activation markers (Extended Data Fig. 5e, f). In addition to confirming the senolytic properties of uPAR-directed CAR T cells, these results indicate that combinatorial strategies using senolytic CAR T cells could be used to treat solid tumours.

Besides cancer, senescence contributes to a range of chronic tissue pathologies, including liver fibrosis—a condition that can evolve into cirrhosis and produces a microenvironment that favours the development of hepatocellular carcinoma[1,2]. As genetic ablation of senescent cells ameliorates liver fibrosis[31,32], we performed dose-escalation studies using m.uPAR-m.28z CAR T cells in the well-defined mouse model of CCl_4-induced liver fibrosis, in which treatment with CCl_4 leads to the accumulation of senescent HSCs, fibrosis and liver damage within six weeks[6]. m.uPAR-m.28z CAR T cells, m.19-m.28z CAR T cells or untransduced T cells were infused at either the previously effective dose of $0.5–1 \times 10^6$ CAR T cells or a higher dosage ($2–3 \times 10^6$) into mice with established liver fibrosis[33] (Extended Data Fig. 6a). In some experiments, mice were treated with m.uPAR-m.28z, m.19-m.28z or control T cells that express click beetle red luciferase to track T cells in vivo using bioluminescence[34] (Extended Data Fig. 6b).

At either dosage, treatment with m.uPAR-m.28z CAR T cells produced a marked reduction in liver fibrosis compared to treatment with m.19-m.28z or untransduced T cells. Hence, liver samples obtained from mice 20 days after treatment with m.uPAR-m.28z CAR T cells had fewer senescent cells and less fibrosis (as assessed by SA-β-gal and Sirius red staining) than controls ($P < 0.001$), and this was associated with an accumulation of adoptively transferred T cells (Fig. 3c, Extended Data Fig. 6c, d). Consistent with on-target activity and a therapeutic benefit, mice that were treated with m.uPAR-m.28z CAR T cells showed reduced serum levels of suPAR and of the liver enzymes alanine aminotransferase (ALT) and aspartate aminotransferase (AST) (Fig. 3d–f, Extended Data Fig. 6e–g), indicating efficient elimination of pro-inflammatory senescent HSCs[31,32] and a reduction in liver damage, respectively. Bioluminescence imaging revealed that transferred T cells at first transited through the lungs as expected. Eventually, uPAR-specific CAR T cells—but not CD19-directed CAR T cells or untransduced T cells—accumulated in the livers of CCl_4-treated mice, showing expansion over a few days followed by rapid contraction (Fig. 3g, h). The high senolytic activity of uPAR CAR T cells was corroborated by an efficient reduction of fibrosis under the aggravated conditions produced by prolonged exposure to CCl_4, as well as a sustained resolution of fibrosis in long-term follow-up studies (Extended Data Fig. 6h, i).

Mice treated at the lower effective dose remained highly active and did not display observable signs of morbidity, changes in temperature or weight or relevant alterations in cell blood counts (Extended Data Fig. 7a–c, e). A moderate infiltration of macrophages was noted in the lungs after 20 days, which also occurred in mice treated with m.19-m.28z CAR T cells or untransduced T cells (Extended Data Fig. 7d). Mice treated at the supratherapeutic dose presented with hypothermia and weight loss, which was accompanied by a rise in serum cytokines including IL-6, GM-CSF, G-CSF and IFNγ (Extended Data Fig. 8a–e). Similar to CAR T cell-associated cytokine-release syndrome (CRS)[35,36], this early toxicity was transient, associated with local accumulation and activation of macrophages and could be mitigated by lower doses of CAR T cells or treatment with CRS-preventing inhibitors of IL-6R and IL-1R (Extended Data

Figs. 8f–i, 9). Altogether, these findings indicate that uPAR-directed CAR T cells at an appropriate dosage can deplete senescent cells without inducing severe CRS-like symptoms, and highlight the potential of short-acting CD28- and CD3ζ-based CAR T cells[37] in senescence-associated indications.

We also tested whether CAR T cells that target uPAR could be effective against fibrosis induced by non-alcoholic steatohepatitis (NASH)—a condition that is increasing in incidence and for which effective therapeutic options are lacking[38]. Although the contribution of cellular senescence to the pathology of NASH is poorly understood, its role in other fibrosis settings prompted us to test two well-established mouse models of NASH for the presence of senescent cells. Indeed, senescent cells were prevalent around the fibrotic areas (Fig. 4a, Extended Data Fig. 10a) and co-expressed uPAR together with either a marker of HSCs (desmin) or a marker of macrophages (F4/80) (Fig. 4b). Accordingly, treatment of mice with diet-induced NASH using 0.5×10^6 m.uPAR-m.28z CAR T cells—but not untransduced control T cells—efficiently eliminated senescent cells, reduced fibrosis and improved liver function (as assessed by serum albumin levels) without eliciting detectable toxicity (Fig. 4c, d, Extended Data Fig. 10b–f). Thus, senolytic CAR T cells are effective against liver fibrosis of different aetiologies.

Perspectives

Here we identify uPAR as a protein that is broadly induced on the surface of senescent cells, and we show that uPAR-targeted CAR T cells can eliminate senescent cells in vitro and in vivo. Owing to its secretion, suPAR serves as a plasma biomarker to assess the senolytic activity of CAR T cells in vivo. Whereas a previous report investigated uPAR as a CAR target in ovarian cancer[39], our results provide proof-of-principle of the therapeutic potential of senolytic CAR T cells in senescence-associated pathologies. Although further work is needed to determine whether uPAR-targeting CAR T cells have the required safety profile to be developed clinically, appropriately dosed senolytic CAR T cells can infiltrate the areas of senescence, efficiently target senescent cells and produce a therapeutic benefit without notable toxicity in mice. Future iterations of this approach could target other cell-surface molecules that are specific to particular senescence contexts, incorporate safety switches[40,41] or use combinatorial strategies to maximize efficacy while minimizing side-effects[14,41].

Unlike tumour cells, senescent cells do not divide or create an immunosuppressive microenvironment, and may present fewer barriers to the development of therapeutically efficacious CAR T cells[18,42]. Furthermore, the rapid waning of senolytic CAR T cells used in our studies soon after their therapeutic action may prove an attractive feature in reducing their potential interference with beneficial aspects of senescence and enabling readministration at later times. Beyond fibrosis, senescence has been linked to many disorders of chronic tissue damage that are associated with ageing, such as severe atherosclerosis, diabetes and osteoarthritis[10]. Consequently, senolytic CAR T cells may have broad therapeutic potential.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2403-9.

1. He, S. & Sharpless, N. E. Senescence in health and disease. *Cell* **169**, 1000–1011 (2017).
2. Sharpless, N. E. & Sherr, C. J. Forging a signature of in vivo senescence. *Nat. Rev. Cancer* **15**, 397–408 (2015).
3. Serrano, M., Lin, A. W., McCurrach, M. E., Beach, D. & Lowe, S. W. Oncogenic *ras* provokes premature cell senescence associated with accumulation of p53 and p16^{INK4a}. *Cell* **88**, 593–602 (1997).
4. Kang, T. W. et al. Senescence surveillance of pre-malignant hepatocytes limits liver cancer development. *Nature* **479**, 547–551 (2011).
5. Demaria, M. et al. An essential role for senescent cells in optimal wound healing through secretion of PDGF-AA. *Dev. Cell* **31**, 722–733 (2014).
6. Krizhanovsky, V. et al. Senescence of activated stellate cells limits liver fibrosis. *Cell* **134**, 657–667 (2008).
7. Collado, M., Blasco, M. A. & Serrano, M. Cellular senescence in cancer and aging. *Cell* **130**, 223–233 (2007).
8. Baker, D. J. et al. Clearance of p16^{Ink4a}-positive senescent cells delays ageing-associated disorders. *Nature* **479**, 232–236 (2011).
9. Baar, M. P. et al. Targeted apoptosis of senescent cells restores tissue homeostasis in response to chemotoxicity and aging. *Cell* **169**, 132–147 (2017).
10. Childs, B. G. et al. Senescent intimal foam cells are deleterious at all stages of atherosclerosis. *Science* **354**, 472–477 (2016).
11. Smith, H. W. & Marshall, C. J. Regulation of cell signalling by uPAR. *Nat. Rev. Mol. Cell Biol.* **11**, 23–36 (2010).
12. Kirkland, J. L. & Tchkonia, T. Cellular senescence: a translational perspective. *EBioMedicine* **21**, 21–28 (2017).
13. Xu, M. et al. Senolytics improve physical function and increase lifespan in old age. *Nat. Med.* **24**, 1246–1256 (2018).
14. Sadelain, M., Rivière, I. & Riddell, S. Therapeutic T cell engineering. *Nature* **545**, 423–431 (2017).
15. Park, J. H. et al. Long-term follow-up of CD19 CAR therapy in acute lymphoblastic leukemia. *N. Engl. J. Med.* **378**, 449–459 (2018).
16. Aghajanian, H. et al. Targeting cardiac fibrosis with engineered T cells. *Nature* (2019).
17. Du, H. et al. Antitumor responses in the absence of toxicity in solid tumors by targeting B7-H3 via chimeric antigen receptor T cells. *Cancer Cell* **35**, 221–237 (2019).
18. Pellegatta, S. et al. Constitutive and TNFα-inducible expression of chondroitin sulfate proteoglycan 4 in glioblastoma and neurospheres: implications for CAR-T cell therapy. *Sci. Transl. Med.* **10**, eaao2731 (2018).
19. Ruscetti, M. et al. NK cell-mediated cytotoxicity contributes to tumor control by a cytostatic drug combination. *Science* **362**, 1416–1422 (2018).
20. Perna, F. et al. Integrating proteomics and transcriptomics for systematic combinatorial chimeric antigen receptor therapy of AML. *Cancer Cell* **32**, 506–519 (2017).
21. Tasdemir, N. et al. BRD4 connects enhancer remodeling to senescence immune surveillance. *Cancer Discov.* **6**, 612–629 (2016).
22. Simon, D. I. et al. Mac-1 (CD11b/CD18) and the urokinase receptor (CD87) form a functional unit on monocytic cells. *Blood* **88**, 3185–3194 (1996).
23. Bugge, T. H. et al. The receptor for urokinase-type plasminogen activator is not essential for mouse development or fertility. *J. Biol. Chem.* **270**, 16886–16894 (1995).
24. Coppé, J. P. et al. Senescence-associated secretory phenotypes reveal cell-nonautonomous functions of oncogenic RAS and the p53 tumor suppressor. *PLoS Biol.* **6**, 2853–2868 (2008).
25. Hayek, S. S. et al. Soluble urokinase receptor and chronic kidney disease. *N. Engl. J. Med.* **373**, 1916–1925 (2015).
26. Belcher, C., Fawthrop, F., Bunning, R. & Doherty, M. Plasminogen activators and their inhibitors in synovial fluids from normal, osteoarthritis, and rheumatoid arthritis knees. *Ann. Rheum. Dis.* **55**, 230–236 (1996).
27. Guthoff, M. et al. Soluble urokinase receptor (suPAR) predicts microalbuminuria in patients at risk for type 2 diabetes mellitus. *Sci. Rep.* **7**, 40627 (2017).
28. Schuliga, M. et al. The fibrogenic actions of lung fibroblast-derived urokinase: a potential drug target in IPF. *Sci. Rep.* **7**, 41770 (2017).
29. Brentjens, R. J. et al. Eradication of systemic B-cell tumors by genetically targeted human T lymphocytes co-stimulated by CD80 and interleukin-15. *Nat. Med.* **9**, 279–286 (2003).
30. Wang, C. et al. Inducing and exploiting vulnerabilities for the treatment of liver cancer. *Nature* **574**, 268–272 (2019).
31. Schnabl, B., Purbeck, C. A., Choi, Y. H., Hagedorn, C. H. & Brenner, D. Replicative senescence of activated human hepatic stellate cells is accompanied by a pronounced inflammatory but less fibrogenic phenotype. *Hepatology* **37**, 653–664 (2003).
32. Puche, J. E. et al. A novel murine model to deplete hepatic stellate cells uncovers their role in amplifying liver damage in mice. *Hepatology* **57**, 339–350 (2013).
33. Kuhn, N. F. et al. CD40 ligand-modified chimeric antigen receptor T cells enhance antitumor function by eliciting an endogenous antitumor response. *Cancer Cell* **35**, 473–488 (2019).
34. Dobrenkov, K. et al. Monitoring the efficacy of adoptively transferred prostate cancer-targeted human T lymphocytes with PET and bioluminescence imaging. *J. Nucl. Med.* **49**, 1162–1170 (2008).
35. Giavridis, T. et al. CAR T cell-induced cytokine release syndrome is mediated by macrophages and abated by IL-1 blockade. *Nat. Med.* **24**, 731–738 (2018).
36. Norelli, M. et al. Monocyte-derived IL-1 and IL-6 are differentially required for cytokine-release syndrome and neurotoxicity due to CAR T cells. *Nat. Med.* **24**, 739–748 (2018).
37. Feucht, J. et al. Calibration of CAR activation potential directs alternative T cell fates and therapeutic potency. *Nat. Med.* **25**, 82–88 (2019).
38. Brunt, E. M. et al. Nonalcoholic fatty liver disease. *Nat. Rev. Dis. Primers* **1**, 15080 (2015).
39. Wang, L. et al. Basing on uPAR-binding fragment to design chimeric antigen receptors triggers antitumor efficacy against uPAR expressing ovarian cancer cells. *Biomed. Pharmacother.* **117**, 109173 (2019).
40. Paszkiewicz, P. J. et al. Targeted antibody-mediated depletion of murine CD19 CAR T cells permanently reverses B cell aplasia. *J. Clin. Invest.* **126**, 4262–4272 (2016).
41. Gargett, T. & Brown, M. P. The inducible caspase-9 suicide gene system as a "safety switch" to limit on-target, off-tumor toxicities of chimeric antigen receptor T cells. *Front. Pharmacol.* **5**, 235 (2014).
42. Anderson, K. G., Stromnes, I. M. & Greenberg, P. D. Obstacles posed by the tumor microenvironment to T cell activity: a case for synergistic therapies. *Cancer Cell* **31**, 311–325 (2017).

Methods

RNA extraction, RNA-seq library preparation and sequencing

Total RNA was isolated from three different models of senescence. (1) *Kras*^{G12D};*p53*$^{-/-}$ cells after 8 days of treatment with vehicle (dimethyl sulfoxide (DMSO)) or combined treatment with the MEK inhibitor trametinib (25 nM) and the CDK4/6 inhibitor palbociclib (500 nM). (2) Oncogene-induced senescent hepatocytes generated in C57BL/6 mice by HTVI. For each mouse, 25 μg of pT3-EF1a-Nras^{G12V}-IRES-GFP-P2A-luciferase plasmid (or pT3-EF1a-Nras$^{G12V;D38A}$-IRES-GFP-P2A-luciferase plasmid as control) and 5 μg CMV-SB13 were suspended in saline solution at the volume of 10% of mouse body weight for administration. Six days after HTVI, mice were anaesthetized and placed on the platform for liver perfusion. Sequential perfusions of Hank's balanced salt solution (HBSS) containing EGTA and HBSS containing collagenase IV were performed, followed by passing the dissociated liver cells through a 100-μm cell strainer. The hepatocytes were then washed again using low-glucose Dulbecco's modified Eagle's medium (DMEM) and centrifuged at a low speed. DAPI-negative and GFP-positive hepatocytes, indicating successful transduction of mutant *Nras* expression, were isolated through low-pressure fluorescence-activated cell sorting (FACS). 3) Senescent or proliferating HSCs (datasets were obtained from a previous study[43]) and proliferating, quiescent or senescent IMR-90 cells (datasets were obtained from a previous study[21]). Sequencing and library preparation were performed at the Integrated Genomics Operation (IGO) at the Memorial Sloan Kettering Cancer Center (MSKCC). RNA-seq libraries were prepared from total RNA. After RiboGreen quantification and quality control by Agilent BioAnalyzer, 100–500 ng of total RNA underwent poly(A) selection and TruSeq library preparation according to the instructions provided by Illumina (TruSeq Stranded mRNA LT Kit, RS-122-2102), with 8 cycles of PCR. Samples were barcoded and run on a HiSeq 4000 or HiSeq 2500 in a 50 bp–50 bp paired-end run, using the HiSeq 3000/4000 SBS Kit or TruSeq SBS Kit v.4 (Illumina) at MSKCC's IGO core facility.

RNA-seq read mapping, differential gene expression analysis and heat map visualization

The resulting RNA-seq data were analysed by removing adaptor sequences using Trimmomatic[44]. RNA-seq reads were then aligned to GRCm38.91 (mm10) with STAR[45] and the transcript count was quantified using featureCounts[46] to generate a raw-count matrix. Differential gene expression analysis and adjustment for multiple comparisons were performed using the DESeq2 package[47] between experimental conditions, with two independent biological replicates per condition, implemented in R (http://cran.r-project.org/). Genes were determined to be differentially expressed on the basis of a greater than two-fold change in gene expression with an adjusted P value of less than 0.05. For heat map visualization of differentially expressed genes, samples were normalized by z-score and plotted using the pheatmap package in R. Transcripts encoding molecules that were determined to be located in the plasma membrane with a confidence score higher than 3 (range 0–5) as determined by UniProtKB were considered cell-surface molecules.

Functional annotations of gene clusters

Pathway enrichment analysis was performed in the resulting gene clusters with the Reactome database using Enrichr[48]. The significance of the tests was assessed using a combined score, described as $c = \log(p) \times z$, in which c is the combined score, p is the Fisher's exact test P value and z is the z-score for deviation from expected rank.

Cell lines and compounds

The following cell lines were used in this study: mouse *Kras*$^{G12D/+}$;*Trp53*$^{-/-}$ (KP) lung cancer cells (provided by T. Jacks and expressing luciferase–GFP as described[19]), and NALM6 and Eμ-ALL01 cells expressing firefly luciferase–GFP[37]. Cells were maintained in a humidified incubator at 37 °C with 5% CO_2. KP cells were grown in DMEM supplemented with 10% fetal bovine serum (FBS) and 100 IU ml^{-1} penicillin–streptomycin. NALM6 and Eμ-ALL01 cells were grown in complete medium composed of RPMI supplemented with 10% FBS, 1% L-glutamine, 1% MEM non-essential amino acids, 1% HEPES buffer, 1% sodium pyruvate, 0.1% β-mercaptoethanol and 100 UI ml^{-1} penicillin–streptomycin. Human primary melanocytes were grown in dermal cell basal medium (ATCC, 200-030) supplemented with the adult melanocyte growth kit (ATCC, 200-042), 10% FBS and 100 IU ml^{-1} penicillin–streptomycin. All cell lines used were negative for mycoplasma.

For drug-induced senescence experiments in vitro, trametinib (S2673) and palbociclib (S1116) were purchased from Selleck Chemicals and dissolved in DMSO to yield 10 mM stock solutions, which were stored at −80 °C[19]. Cells were treated with MEK inhibitor (25 nM) and CDK4/6 inhibitor (500 nM). The growth medium was changed every two days. For in vivo experiments trametinib was dissolved in a 5% hydroxypropyl methylcellulose and 2% Tween-80 solution (Sigma) and palbociclib was dissolved in sodium lactate buffer (pH 4) (as described previously[19]). Mice were treated with 1 mg per kg body weight of trametinib and 150 mg per kg body weight of palbociclib as previously described[19]. Caerulein was purchased from Bachem. Anakinra was purchased from Sobi and administered intraperitoneally at a dose of 30 mg per kg body weight twice a day for 8 days starting 24 h before transfer of CAR T cells. Anti-mouse IL-6R (clone MP5-20F3) was purchased from BioXCell and administered intraperitoneally once per day at 25 mg per kg body weight for the first dose and 12.5 mg per kg body weight for subsequent doses for 8 days starting 24 h before transfer of CAR T cells as previously described[35].

SA-β-gal staining

SA-β-gal staining was performed as previously described[19] at pH 6.0 for human cells and tissue and pH 5.5 for mouse cells and tissue. Fresh frozen tissue sections or adherent cells plated in 6-well plates were fixed with 0.5% glutaraldehyde in phosphate-buffered saline (PBS) for 15 min, washed with PBS supplemented with 1 mM $MgCl_2$ and stained for 5–8 h in PBS containing 1 mM $MgCl_2$, 1 mg ml^{-1} X-gal, 5 mM potassium ferricyanide and 5 mM potassium ferrocyanide. Tissue sections were counterstained with eosin. Five high power fields per well or section were counted and averaged to quantify the percentage of SA-β-gal^{+} cells.

Quantitative PCR with reverse transcription

Total RNA was isolated using the RNeasy Mini Kit (Qiagen) and cDNA was obtained using TaqMan reverse-transcription reagents (Applied Biosystems). Quantitative PCR (qPCR) was performed in triplicates using SYBR green PCR master mix (Applied Biosystems) on the ViiA 7 Real-Time PCR System (Invitrogen). *GAPDH* or *ACTB* served as endogenous normalization controls for mouse and human samples.

Mice

All mouse experiments were approved by the MSKCC Internal Animal Care and Use Committee. All relevant animal use guidelines and ethical regulations were followed. Mice were maintained under specific pathogen-free conditions, and food and water were provided ad libitum. The following mice were used: C57BL/6N background, NOD-scid IL2Rg$^{\text{null}}$ (NSG) mice (purchased from The Jackson laboratory) and B6.SJL-Ptrc^{a}/BoyAiTac (CD45.1 mice) (purchased from Taconic). Mice of both sexes were used at 8–12 weeks of age (5–7 weeks old for the xenograft experiments and 6–10 weeks old for T cell isolation) and were kept in group housing. Mice were randomly assigned to the experimental groups.

Transposon-mediated intrahepatic gene transfer

Transposon-mediated intrahepatic gene transfer was performed as previously described[4]. In brief, 8–12-week-old C57BL/6J mice received a saline solution at a final volume of 10% of their body weight containing 30 μg of total DNA composed of a 5:1 molar ratio of transposon-encoding vector (containing either the sequence for *Nras*^{G12V} or the sequence for the GTPase-dead form *Nras*$^{G12V/D38A}$) to transposase-encoding vector (*Sleeping Beauty* 13) through HTVI. For CAR T cell studies, NSG mice were intravenously injected with 0.5×10^6 human CAR T cells or untransduced T cells 10 days after HTVI and monitored by bioluminescence imaging using the IVIS Imaging System (PerkinElmer) with Living Image software (PerkinElmer). At day 15 after CAR injection, mice were euthanized and livers were removed and used for further analysis.

Generation of mouse pancreatic intraepithelial neoplasias

The mouse strain has been previously described[49]. To induce pancreatic intraepithelial neoplasias, KC;RIK (p48-Cre;RIK;LSLKrasG12D) male mice were treated with eight (one per hour) intraperitoneal injections of 80 µg kg^{-1} caerulein (Bachem) for two consecutive days. Mice were then euthanized 21 weeks later and their pancreases were used for further analysis. Age-matched C;RIK mice (expressing wild-type *Kras*) injected with PBS were used as controls for normal pancreas.

In vivo induction of CCl_4-induced liver fibrosis

C57BL/6N mice were treated twice a week with 12 consecutive intraperitoneal injections of 1 ml kg^{-1} tetrachloride (CCl_4) to induce liver fibrosis[6,43]. For mouse CAR T cell studies, cyclophosphamide (200 mg kg^{-1}) was administered 16–24 h before T cell injection. Mice received $0.5–1 \times 10^6$ or $2–3 \times 10^6$ CAR T cells or untransduced T cells (same total numbers of T cells) and CCl_4 was continuously administered at the same dose and interval until day 20 after CAR T cell injection, when mice were euthanized 48–72 h after the last CCl_4 injection. Blood was collected by facial vein puncture or cardiac puncture.

In vivo induction of NASH-induced liver fibrosis

C57BL/6N mice were fed with a NASH-inducing diet (Teklad TD.160785, which contains 10.2% kcal from protein, 37.3% kcal from carbohydrate and 52.6% kcal from fat) and fructose-containing drinking water (23.1 g fructose and 18.9 g glucose dissolved in 1 l water and then filter-sterilized) at 8–10 weeks of age[50,51]. Body weight was measured weekly. For mouse CAR T cell studies, cyclophosphamide (200 mg kg^{-1}) was administered 16 h before T cell injection. Mice received 0.5×10^6 CAR T cells or untransduced T cells (same total numbers of T cells) and they received the same NASH diet until day 20 after T cell injection, when they were euthanized. Blood was collected by facial vein puncture or cardiac puncture.

For the 'STAM' model[52], liver tissue samples (unstained slides for immunohistochemistry) were purchased from SMC laboratories.

Patient-derived xenografts

Experiments with patient-derived xenografts were performed as described[19], using 5–7-week-old female NSG mice. MSK-LX27 was derived from a lung adenocarcinoma containing *KRAS*^{G12D} and *P53* (*P53* is also known as *TP53*) mutations and a deletion in *CDKN2A* and was cut into pieces and inserted in the subcutaneous space. Mice were monitored daily, weighed twice weekly and caliper measurements began when tumours became visible. Tumours were measured using the formula: tumour volume = $(D \times d^2)/2$ (in which D is the longer diameter and d is the shorter diameter) and when they reached a size of 100–200 mm^3, mice were randomized on the basis of the starting tumour volume and treated with vehicle or trametinib (3 mg per kg body weight) and palbociclib (150 mg per kg body weight) orally for 4 consecutive days followed by 3 days off treatment. Experimental end points were achieved when tumours reached a size of 2,000 mm^3 or became ulcerated. Tumours were collected at the experimental end point and tissue was divided evenly for 10% formalin fixation and optimal cutting temperature (OCT) compound frozen blocks.

Patient samples

De-identified human samples from liver biopsies of patients with liver fibrosis from viral (hepatitis B or C), alcoholic and non-alcoholic fatty liver disease were obtained through the Department of Pathology at Mount Sinai Hospital. Human pancreatic intraepithelial neoplasia samples were obtained through the Department of Pathology at MSKCC. Human atherosclerosis samples were obtained through the Department of Pathology at Weill Medical College of Cornell University. All human studies complied with all relevant guidelines and ethical regulations and were approved by the Institutional Review Board at Mount Sinai, Weill Medical College or MSKCC.

Histological analysis

Tissues were fixed overnight in 10% formalin, embedded in paraffin and cut into 5-µm sections. Sections were subjected to haematoxylin and eosin (H&E) staining, and to Sirius red staining for fibrosis detection. For fibrosis quantification, at least three whole sections from each mouse were scanned and the images were quantified using NIH ImageJ software. The amount of fibrotic tissue was calculated relative to the total analysed liver area as previously described. Immunohistochemical and immunofluorescence staining was performed following standard protocols. The following primary antibodies were used: anti-human uPAR (R&D, AF807, lot BBS0318071, 1:50), anti-mouse uPAR (R&D, AF534, lot DCL0418021, 1:50), anti-mouse NRAS (Santa Cruz, SC-31, lot A1020, 1:50), anti-mouse SMA (Abcam, Ab5694, lot GR283004-16, 1:50), anti-mouse KATE (Evrogen, ab233, lot 23301201267, 1:1,000), anti-human CD3 (Abcam, ab5690, lot GR3220039-4, 1:100), Myc-tag (Cell Signaling, 2276S, lot 24, 1:50), anti-mouse Ki-67 (Abcam, ab16667, lot GR3305281-1, 1:200), anti-mouse IL-6 (Abcam, ab6672, lot GR3195128-19, 1:50), p16-INK4A (Proteintech, 10883-1-AP, lot 00057396, 1:50), anti-mouse P-ERKT202/Y204 (Cell Signaling.4370, lot 1:800), desmin (Thermo Fisher Scientific, RB-9014, lot 9014p1806Q, 1:200), AF488 donkey anti-rabbit (Invitrogen, A21206, lot 1874771, 1:500), AF488 donkey anti-mouse (Invitrogen, A21202, lot 1820538, 1:500), AF594 donkey anti-rabbit (Invitrogen, A21207, lot 1602780, 1:500), AF594 donkey anti-mouse (Invitrogen, A21203, lot 1163390, 1:500), AF594 donkey anti-goat (Invitrogen, A11058, lot 2045324, 1:500), AF594 goat anti-rat (Invitrogen, A11007, lot 1903506, 1:500).

Flow cytometry

For analysis of uPAR expression in cell lines after induction of senescence, KP cells were treated with trametinib (25 nM) and palbociclib (500 nM) or with vehicle (DMSO), and human primary melanocytes were continuously passaged for 15 passages and then trypsinized, resuspended in PBS supplemented with 2% FBS and stained with the following antibodies for 30 min on ice: PE-conjugated anti-mouse uPAR (R&D, FAB531P) or APC-conjugated anti-human uPAR (Thermo Fisher Scientific, 17-3879-42). The following fluorophore-conjugated antibodies were used for in vitro and in vivo experiments in the indicated dilutions ('h' prefix denotes anti-human; 'm' prefix denotes anti-mouse): hCD45 APC-Cy7 (clone 2D1, BD, 557833, lot 9081815, 1:100), hCD4 BUV395 (clone SK3, BD, 563550, lot 6252529, 1:100), hCD4 BV480 (clone SK3, BD, 566104, lot 8092993, 1:50), hCD62L BV421 (clone DREG-56, BD, 563862, lot 8194954, 1:100), hCD45RA BV650 (clone HI100, BD, 563963, lot 9057952, 1:100), hPD-1 BV480 (clone EH12.1, BD, 566112, lot 8235507, 1:100), hCD19 BUV737 (clone SJ25C1, BD, 564303, lot 8130572, 1:100), hCD271 PE (clone C40-1457, BD, 557196, lot 7068641, 1:100), hIL-2 PE-Cy7 (clone MQ1-17H12, Invitrogen, 25-7029-42, lot 4336863, 1:50), hTNF BV650 (clone Mab11, BD, 563418, lot 7082880, 1:50), hIFNγ BUV395 (clone B27, BD, 563563, lot 6320836, 1:50) hTIM3 BV785 (clone F38-2E2, Biolegend, 345032, lot B265346, 1:100), hCD8 PE-Cy7 (clone SK1, eBioscience, 25-0087-42, lot 2066348, 1:100), hCD8 APC-Cy7 (clone SK1, BD, 557834, lot 7110951, 1:50), hCD223 PerCP-eFluor710 (clone 3DS223H, eBioscience, 46-2239-42, lot 4321735, 1:100), hGrB APC (clone GB12, Invitrogen, MHGB05, lot 1884625, 1:67), hMyc-tag AF647 (clone 9B11, Cell Signaling Technology, 2233S, lot 23, 1:50), hCD19 PB (clone SJ25-C1, Invitrogen, MHCD1928, 1:100), hCD87 APC (clone VIM5, eBioscience, 17-3879-42, lot 17-3879-42, 1:50), hCD87 PerCp-eFluor710 (clone VIM5, eBioscience, 46-3879-42, lot 46-2239-42, 1:50), muPAR PE (R&D Systems, FAB531P, lot ABLH0419081, 1:50), muPAR AF700 (R&D Systems, FAB531N, lot 1552229, 1:50), mCD45.1 APC-Cy7 (clone A20, Biolegend, 110716, lot B285685, 1:200), m.CD45.1 BV785 (clone A20, Biolegend, 110743, lot B270183, 1:100), mCD45.2 PE (clone 104, Biolegend, 109808, lot B271929, 1:100), mCD45.2 AF700 (clone 104, Biolegend, 109822, lot B252126, 1:200), mSiglec-F PerCP-Cy5.5 (clone E50-2440, BD, 565526, lot 8232650, 1:200), mI-A/I-E BV605 (clone M5/114.15.2, Biolegend, 107639, lot B293222, 1:50), mF4/80 BV421 (clone T45-2342, BD, 565411, lot 8330526, 1:200), mCD11b BUV395 (clone M1/70, BD, 563553, lot 8339988, 1:200), mCD11c BV650 (clone N418, Biolegend, 117339, lot B253523, 1:200), mLY6G BV510 (clone 1A8, Biolegend, 127633, lot B266675, 1:200), mLY6G APC/Fire750 (clone 1A8, Biolegend, 127652, lot B274284, 1:100), miNOS PE-Cy7 (clone CXNFT, eBioscience, 25-5920-82, lot 2127491, 1:200), mCD19 PE (clone 1D3/CD19, Biolegend, 152408, lot B260181, 1:100), mCD25 BV605 (clone PC61, Biolegend, 102035, lot B291215, 1:50), mCD69 PerCpCy5.5 (clone H1.2F3, Biolegend, 104522, lot B244018, 1:100), mCD3 AF488 (clone 17A2, Biolegend, 100210, lot B284975, 1:100), mCD4 BUV395 (clone GK1.5, BD, 563790, lot 9101822, 1:50), mCD4 FITC (clone GK1.5, BD, 553729, lot 9204449, 1:50) and mCD8 PE-Cy7 (Clone: 53-6.7, Biolegend, 100722, lot B282418, 1:50). 7-AAD (BD, 559925, lot 9031655, 1:40), DAPI (Life Technologies D1306), Fixable Viability Dye eFluor 506 (65-0866-14, eBioscience, lot 2095423, 1:200) and LIVE/DEAD Fixable Violet (L34963, Invitrogen, lot 1985351, 1:100) were used as a viability dyes.

CAR staining was performed with Alexa Fluor 647 AffiniPure F(ab′)$_2$ Fragment Goat Anti-Rat IgG (Jackson ImmunoResearch, 112-6606-072). For cell counting, CountBright Absolute Counting Beads were added (Invitrogen) according to the manufacturer's instructions. For in vivo experiments, Fc receptors were blocked using FcR blocking reagent, mouse (Miltenyi Biotec). For intracellular cytokine secretion assay, cells were fixed and permeabilized using the Cytofix/Cytoperm Fixation/Permeabilization Solution Kit (BD Biosciences) or Intracellular Fixation & Permeabilization Buffer Set Kit (eBioscience, 88-8824-00) according to the manufacturer's instructions.

Flow cytometry was performed on a LSRFortessa instrument (BD Biosciences) or Cytek Aurora (CYTEK) and data were analysed using FlowJo (TreeStar).

For in vivo sample preparation, livers were dissociated using the MACS liver dissociation kit (Miltenyi Biotec, 130-1-5-807), filtered through a 100-μm strainer and washed with PBS, and red blood cell lysis was achieved with an ACK (ammonium–chloride–potassium) lysing buffer (Lonza). Cells were washed with PBS, resuspended in FACS buffer and used for subsequent analysis. Lungs were minced and digested with 1mg/ml collagenase type IV and DNase type IV in RPMI at 37C and 200rpm for 45 min, filtered through 100μm strainer, washed with PBS, and red blood cell lysis was achieved with an ACK lysing buffer (Lonza). Cells were washed with PBS, resuspended in FACS buffer and used for subsequent analysis. For bone marrow samples, tibias and femurs were mechanically disrupted with a mortar in PBS and 2 mM EDTA, filtered through a 40-μm strainer and washed with PBS and 2 mM EDTA, and red blood cell lysis was achieved with an ACK lysing buffer (Lonza). Cells were washed with PBS and 2 mM EDTA, resuspended in FACS buffer and used for subsequent analysis. Spleens were mechanically disrupted with the back of a 5-ml syringe, filtered through a 40-μm strainer and washed with PBS and 2 mM EDTA and red blood cell lysis was achieved with an ACK lysing buffer (Lonza). Cells were washed with PBS and 2 mM EDTA, resuspended in FACS buffer and used for subsequent analysis.

Cytokine measurements

Serum cytokines were measured using cytometric bead arrays (BD) as per the manufacturer's instructions.

Detection of suPAR levels

suPAR levels from cell culture supernatant or mouse plasma were evaluated by enzyme-linked immunosorbent assay (ELISA) according to the manufacturer's protocol (R&D systems, DY531 (mouse) or DY807 (human)).

Liver function tests

The levels of ALT, AST and albumin in mouse serum were measured according to the manufacturer's protocol, using the EALT-100 (ALT), EASTR-100 (AST) and DIAG-250 (albumin) kits from BioAssay systems.

Isolation, expansion and transduction of human T cells

All blood samples were handled following the required ethical and safety procedures. Peripheral blood was obtained from healthy volunteers and buffy coats from anonymous healthy donors were purchased from the New York Blood Center. Peripheral blood mononuclear cells were isolated by density gradient centrifugation. T cells were purified using the human Pan T Cell Isolation Kit (Miltenyi Biotec), stimulated with CD3/CD28 T cell activator Dynabeads (Invitrogen) as described[37] and cultured in X-VIVO 15 (Lonza) supplemented with 5% human serum (Gemini Bio-Products), 5 ng ml^{-1} interleukin-7 and 5 ng ml^{-1} interleukin-15 (PeproTech). T cells were counted using an automated cell counter (Nexcelom Bioscience).

Forty-eight hours after initiating T cell activation, T cells were transduced with retroviral supernatants by centrifugation on RetroNectin-coated plates (Takara). Transduction efficiencies were determined four days later by flow cytometry and CAR T cells were adoptively transferred into mice or used for in vitro experiments.

Isolation, expansion and transduction of mouse T cells

B6.SJL-Ptrc^{a}/BoyAiTac mice (CD45.1 mice) were euthanized and spleens were collected. After tissue dissection and red blood cell lysis, primary mouse T cells were purified using the mouse Pan T cell Isolation Kit (Miltenyi Biotec). Purified T cells were cultured in RPMI-1640 (Invitrogen) supplemented with 10% FBS (HyClone), 10 mM HEPES (Invitrogen), 2 mM L-glutamine (Invitrogen), MEM non-essential amino acids 1×(Invitrogen), 55 uM β-mercaptoethanol, 1 mM sodium pyruvate (Invitrogen), 100 IU ml^{-1} recombinant human IL-2 (Proleukin; Novartis) and mouse anti-CD3/28 Dynabeads (Gibco) at a bead:cell ratio of 1:2. T cells were spinoculated with retroviral supernatant collected from Phoenix-ECO cells 24 h after initial T cell activation as described[33,53] and used for functional analysis 3–4 days later.

Genetic modification of T cells

The human and mouse SFG γ-retroviral m.uPAR-28z plasmids were constructed by stepwise Gibson assembly (New England BioLabs) using the SFG-1928z backbone as previously described[29,54–56]. The amino acid sequence for the single-chain variable fragment (scFv) specific for mouse uPAR was obtained from the heavy and light chain variable regions of a selective monoclonal antibody against mouse uPAR (R&D MAB531-100) through mass spectometry performed by Bioinformatics Solutions. In the human SFG-m.uPAR-h.28z CARs, the anti-mouse uPAR scFv is thus preceded by a human CD8A leader peptide and followed by CD28 hinge–transmembrane–intracellular regions, and CD3z intracellular domains linked to a P2A sequence to induce co-expression of truncated LNGFR. In the mouse SFG-m.uPAR-m.28z CARs, the anti-mouse uPAR scFv is preceded by a mouse CD8A leader peptide and followed by the Myc-tag sequence (EQKLISEEDL), mouse CD28 transmembrane and intracellular domain and mouse CD3z intracellular domain[33].

Plasmids encoding the SFGγ retroviral vectors were used to transfect gpg29 fibroblasts (H29) to generate VSV-G pseudotyped retroviral supernatants, which were used to construct stable retrovirus-producing cell lines as described[29,33]. For T cell imaging studies, mouse T cells were transduced with retroviral supernatants encoding SFG-GFP-click beetle red luciferase[57].

Cytotoxicity assays

The cytotoxicity of CAR T cells was determined by standard luciferase-based assays or by calcein-AM-based cytotoxicity assays. For luciferase-based assays, target cells expressing firefly luciferase (FFLuc-GFP) were co-cultured with T cells in triplicate at the indicated effector:target ratios using black-walled 96-well plates with 5×10^4 (for NALM6 and Eμ-ALL01) or 1.5×10^4 (for KP) target cells in a total volume of 100 µl per well in RPMI or DMEM medium, respectively. Target cells alone were plated at the same cell density to determine the maximum luciferase expression (relative light units (RLU)) and maximum release was determined by addition of 0.2% Triton-X100 (Sigma). Either 4 or 18 h later, 100 µl luciferase substrate (Bright-Glo, Promega) was directly added to each well. Emitted light was detected in a luminescence plate reader. Lysis was determined as $(1-(\mathrm{RLU}_{\mathrm{sample}})/(\mathrm{RLU}_{\mathrm{max}})) \times 100$. For calcein-AM-based assays, target cells (NALM6) were loaded with 20 µM calcein-AM (Thermo Fisher Scientific) for 30 min at 37 °C, washed twice and co-incubated with T cells in triplicate at the indicated effector:target ratios in 96-well round-bottomed plates with 5×10^3 target cells in a total volume of 200 µl per well in complete medium. Target cells alone were plated at the same cell density to determine spontaneous release and maximum release was determined by incubating the targets with 0.2% Triton-X100 (Sigma). After a 4-h co-culture, supernatants were collected and free calcein was quantitated using a Spark plate reader (Tecan). Lysis was calculated as: ((experimental release − spontaneous release)/(maximum release − spontaneous release)) × 100.

Statistical analysis and figure preparation

Data are presented as mean ± s.e.m. Statistical analysis was performed by Student's *t*-test using GraphPad Prism v.6.0 or 7.0 (GraphPad software). *P* values of less than 0.05 were considered to be statistically significant. Survival was determined using the Kaplan–Meier method. No statistical methods were used to predetermine sample size in the mouse studies, and mice were allocated at random to treatment groups. The investigators were not blinded to allocation during experiments and outcome assessment. Figures were prepared using BioRender.com for scientific illustrations and Illustrator CC 2019 (Adobe).

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The RNA-seq data have been deposited in the Gene Expression Omnibus under the accession number GSE145642. Source data are provided with this paper. All other data supporting the findings of this study will be made available upon reasonable request to the corresponding authors. Source data

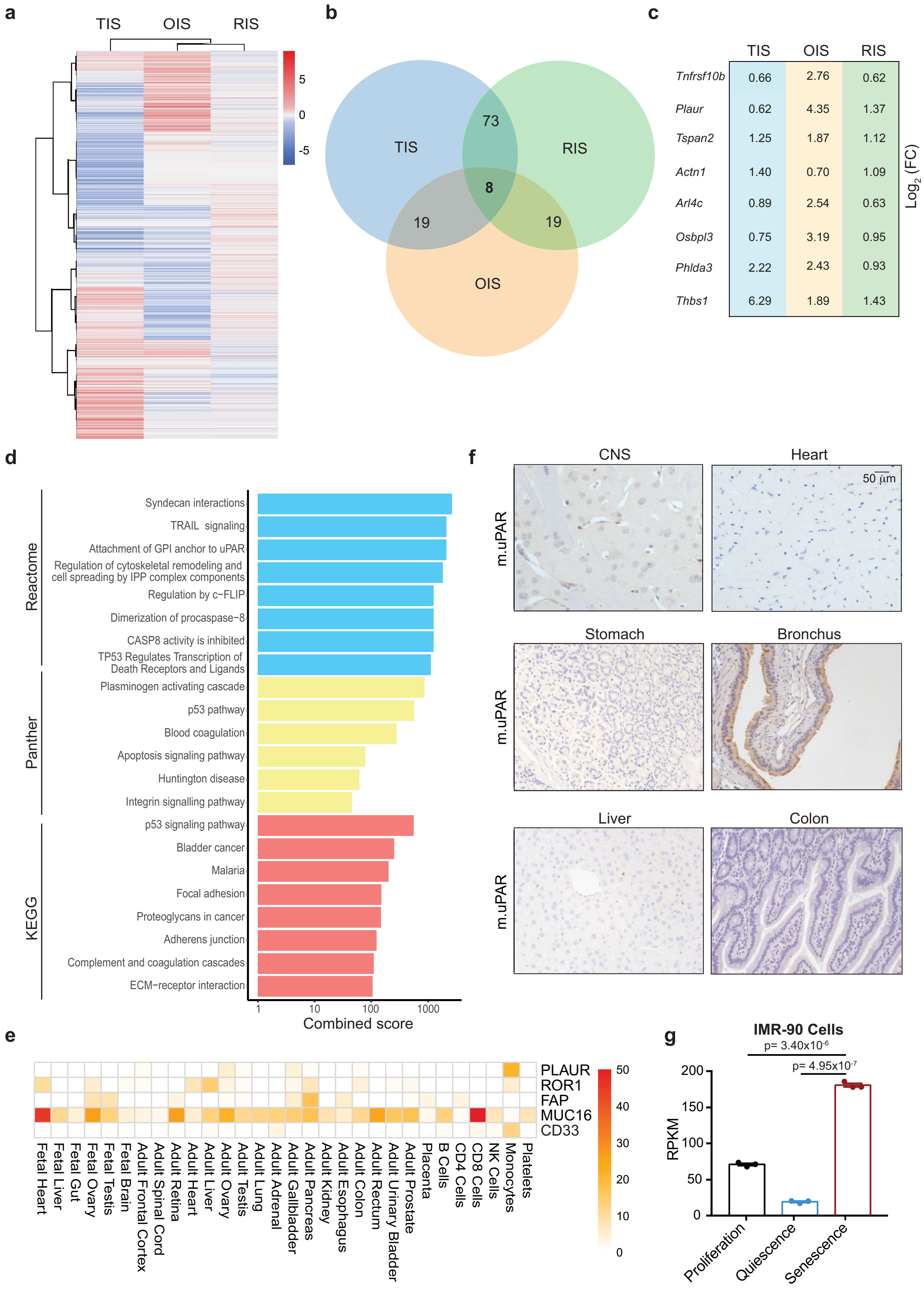
43. Lujambio, A. et al. Non-cell-autonomous tumor suppression by p53. *Cell* **153**, 449–460 (2013).
44. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
45. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
46. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
47. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. **15**, 550 (2014).
48. Chen, E. Y. et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128 (2013).
49. Livshits, G. et al. Arid1a restrains Kras-dependent changes in acinar cell identity. *eLife* **7**, e35216 (2018).
50. Zhu, C. et al. Hepatocyte Notch activation induces liver fibrosis in nonalcoholic steatohepatitis. *Sci. Transl. Med*. **10**, eaat0344 (2018).
51. Wang, X. et al. Hepatocyte TAZ/WWTR1 promotes inflammation and fibrosis in nonalcoholic steatohepatitis. *Cell Metab*. **24**, 848–862 (2016).
52. Fujii, M. et al. A murine model for non-alcoholic steatohepatitis showing evidence of association between diabetes and hepatocellular carcinoma. *Med. Mol. Morphol*. **46**, 141–152 (2013).
53. Davila, M. L., Kloss, C. C., Gunset, G. & Sadelain, M. CD19 CAR-targeted T cells induce long-term remission and B cell aplasia in an immunocompetent mouse model of B cell acute lymphoblastic leukemia. *PLoS ONE* **8**, e61338 (2013).
54. Maher, J., Brentjens, R. J., Gunset, G., Rivière, I. & Sadelain, M. Human T-lymphocyte cytotoxicity and proliferation directed by a single chimeric TCRζ/CD28 receptor. *Nat. Biotechnol*. **20**, 70–75 (2002).
55. Brentjens, R. J. et al. Genetically targeted T cells eradicate systemic acute lymphoblastic leukemia xenografts. *Clin. Cancer Res*. **13**, 5426–5435 (2007).
56. Hagani, A. B., Rivière, I., Tan, C., Krause, A. & Sadelain, M. Activation conditions determine susceptibility of murine primary T-lymphocytes to retroviral infection. *J. Gene Med*. **1**, 341–351 (1999).
57. Santos, E. B. et al. Sensitive in vivo imaging of T cells using a membrane-bound *Gaussia princeps* luciferase. *Nat. Med*. **15**, 338–344 (2009).
58. Van der Schueren, B. et al. Low cytochrome oxidase 4I1 links mitochondrial dysfunction to obesity and type 2 diabetes in humans and mice. *Int. J. Obes*. **39**, 1254–1263 (2015).

Acknowledgements We thank A. Lujambio and R. Brody and the Biorepository and Pathology Core at Icahn School of Medicine at Mount Sinai, and G. Askan and O. Basturk at the Department of Pathology at MSKCC, for tissue samples; L. Zender and H. Chen for sharing plasmids; N. Salgado, H. Chen, T. Baslan, S. Tian, A. Wuest, W. Luan and G. Gunset for technical assistance; and C. J. Sherr, E. de Stanchina, N. Kuhn, A. Dobrin, M. L. Sjöstrand and other members of the Lowe and Sadelain laboratories for insightful discussions. This work was supported by a grant from the National Institute of Aging (AG065396) to S.W.L., the Pasteur-Weizmann/Servier award to M.S. and a Memorial Sloan Kettering Cancer Center support grant (P30 CA008748) to both S.W.L. and M.S. laboratories. S.L.F. was supported by a grant from the National Institute of Diabetes and Digestive and Kidney Diseases (R01DK56621), a grant from the Department of Defense (CA150272) and the P30 grant (CA165979); C.A. was supported by a postgraduate fellowship from La Caixa foundation and is the recipient of the Harold E. Varmus graduate student fellowship from the Gerstner Sloan Kettering graduate school; J.F. was supported by the Care-for-Rare Foundation and the German Research Foundation (DFG); J.L. was supported by a fellowship from the DFG and a Shulamit Katzman Endowed Postdoctoral Research Fellowship; J.F. and J.L. are part of the Experimental Medicine Program at the University of Tuebingen; D.A.-C. was supported by a postdoctoral fellowship from Fundación Ramón Areces; J.A.B. was supported by the Grayer postgraduate fellowship and the Geoffrey Been graduate student fellowship from the Gerstner Sloan Kettering graduate school; and A.K. was supported by a grant from the National Cancer Institute (U54 OD020355-01). S.W.L. is the Geoffrey Been Chair of Cancer Biology and a Howard Hughes Medical Institute Investigator. We thank the following MSKCC core facilities for support: SKI flow cytometry core facility, animal facility, antitumor assessment core, laboratory for comparative biology, bioinformatics core and integrated genomics operation core.

Author contributions C.A., J.F. and J.L. conceived the project, designed, performed and analysed experiments and wrote the paper with assistance from all authors. Y.-J.H. analysed RNA-seq data. C.Z., D.A.-C., J.M.-S., J.A.B., X.L., A.K., S.H. and T.G. performed and analysed experiments. A.P. performed histopathological toxicity analysis. S.L.F. provided human liver samples, analysed data and reviewed the manuscript. E.P. provided human carotid endarterectomy samples and reviewed the manuscript. V.P. helped with T cell imaging studies. M.S. and S.W.L. conceived the project, supervised experiments and wrote the paper. All authors read and approved of the paper.

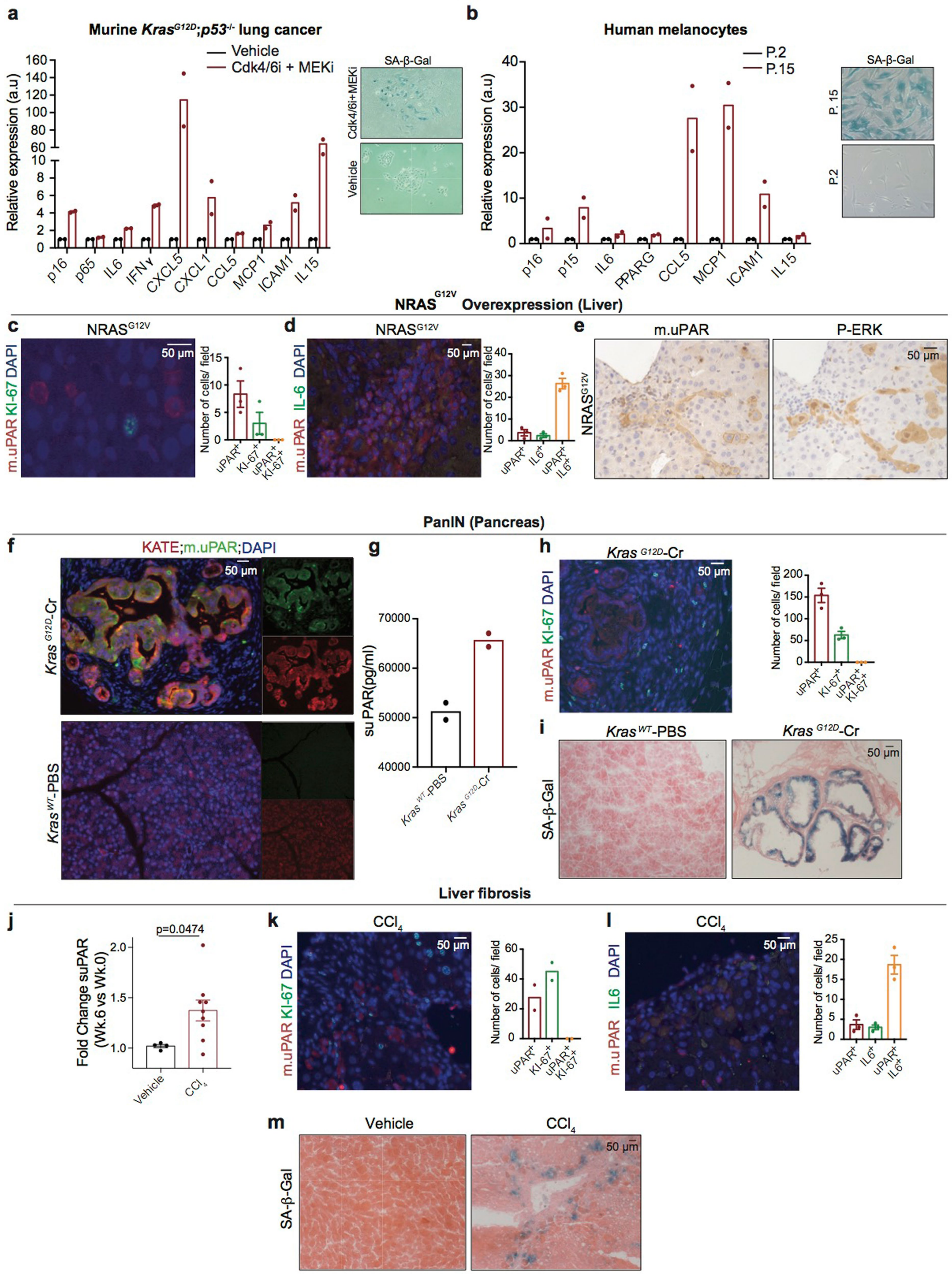
Competing interests A patent application (PCT/US2020/016290; filed on 02/01/2020) has been submitted based in part on results presented in this manuscript on the use of CAR T cells that target uPAR as senolytic agents. C.A., J.F., J.L., M.S. and S.W.L. are listed as the inventors. J.F. and M.S. hold other unrelated patents on CAR technologies.S.W.L. is an advisor for and has equity in the following biotechnology companies: ORIC Pharmaceuticals, Faeth Therapeutics, Blueprint Medicines, Geras Bio, Mirimus Inc., PMV Pharmaceuticals and Constellation Pharmaceuticals.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2403-9.
Correspondence and requests for materials should be addressed to M.S. or S.W.L.
Peer review information *Nature* thanks Jesus Gil, Stephen Gottschalk and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.
Reprints and permissions information is available at http://www.nature.com/reprints.



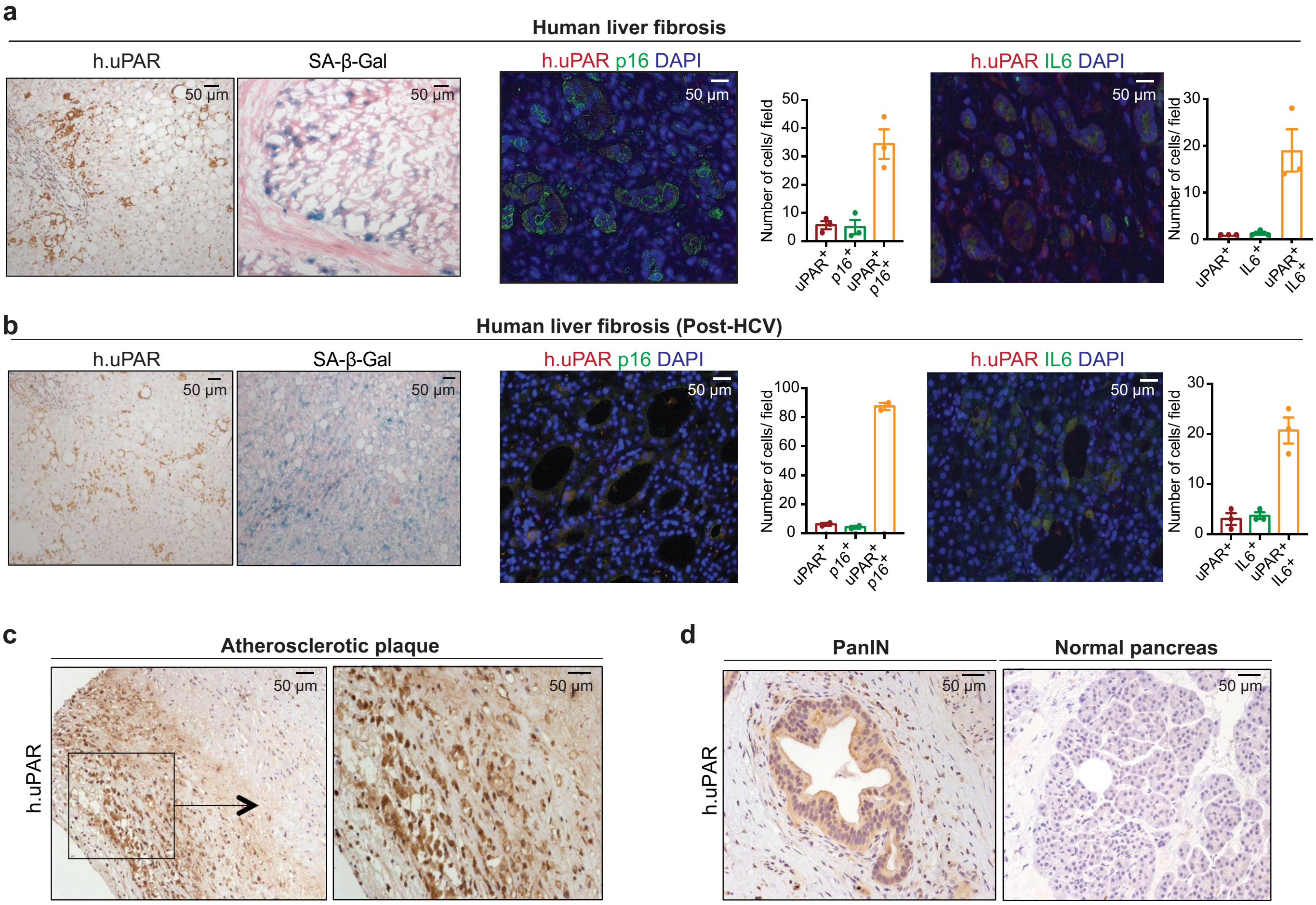
Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Genes encoding surface molecules that are commonly upregulated in senescence. **a**, Heat map of genes upregulated in therapy-induced senescence (TIS), oncogene-induced senescence (OIS) or replication-induced senescence (RIS) in HSCs. **b**, Venn diagram showing the number of common genes upregulated in the three datasets in **a**. **c**, Fold change ($\log_2$(expression in senescent cells/expression in non-senescent cells)) of the eight commonly upregulated genes in the three different datasets in **a**. **d**, Combined enrichment score of significantly enriched gene sets among the eight commonly upregulated genes in senescence. ECM, extracellular matrix; GPI, glycosylphosphatidylinositol. **e**, Heat map showing the expression profile of uPAR (*PLAUR*) in human vital tissues (as determined by the Human Proteome Map) compared to the expression profiles of other targets of CAR T cells in clinical trials. NK cells, natural killer cells. **f**, Immunohistochemical staining of mouse uPAR (m.uPAR) in vital tissues of C57BL/6J mice. Representative results of $n = 2$ independent experiments. **g**, Reads per kilobase (RPKM) of *PLAUR* mRNA in proliferating, quiescent (induced by serum starvation) or senescent (triggered by overexpression of $HRAS^{G12V}$) human IMR-90 fibroblasts. Results of one independent experiment with $n = 3$ replicates for proliferating, quiescent and senescent conditions. Data are mean ± s.e.m.; two-tailed unpaired Student's *t*-test.

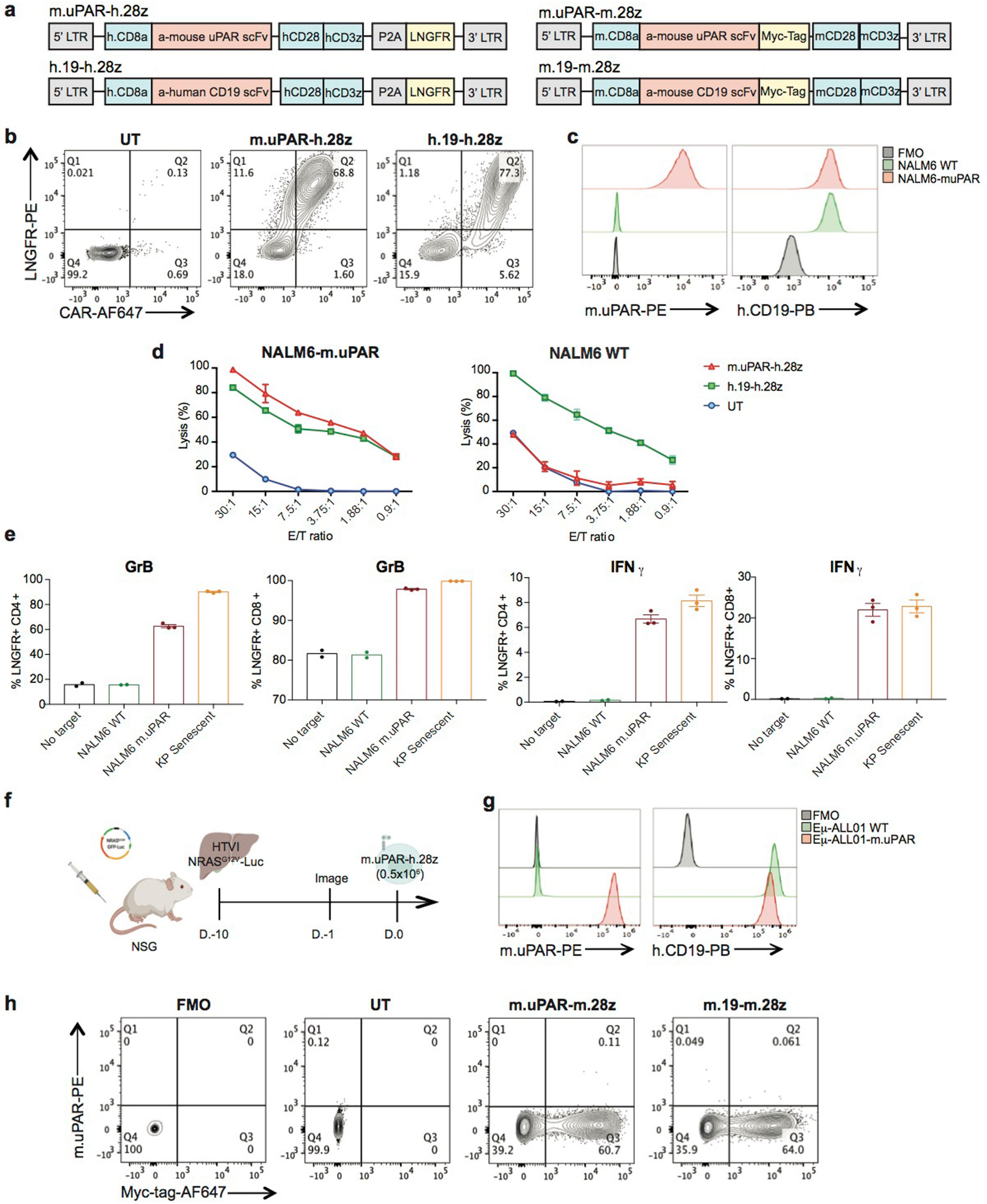


Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | uPAR is a cell-surface and secreted biomarker of senescence. **a**, **b**, qPCR of SASP-associated gene expression in senescent versus proliferating mouse KP tumour cells (**a**) or human primary melanocytes (**b**) and representative SA-β-gal staining; a.u. arbitrary units. **c**, **d**, Co-immunofluorescence staining and quantifications of uPAR (red) and Ki-67 (green) (**c**) or uPAR (red) and IL-6 (green) (**d**). **e**, Immunohistochemical staining of uPAR or phosphorylated ERK (P-ERK) in serial sections of mouse livers six days after transfection by HTVI with a plasmid encoding *Nras*^{G12V}. Representative results of two independent experiments ($n = 3$ mice per group). **f**–**i**, Mice expressing endogenous *Kras*^{G12D} in pancreatic epithelial cells were treated with caerulein (Cr) and euthanized 21 weeks afterwards when they had developed pancreatic intraepithelial neoplasias. Age-matched C;RIK mice (expressing wild-type *Kras*) injected with PBS were used as controls. **f**, Co-immunofluorescence staining of KATE (red) and uPAR (green). Representative results of two independent experiments ($n = 3$ mice per group). **g**, Levels of suPAR in the mice in **f**. Representative results of two independent experiments ($n = 2$ mice per group). **h**, Co-immunofluorescence staining and quantification of uPAR (red) and Ki-67 (green). Representative results of two independent experiments ($n = 3$ mice per group). **i**, Representative SA-β-gal staining. Representative results of one independent experiment ($n = 3$ mice per group). **j**–**m**, Mice were treated with either vehicle or CCl_4 twice weekly for six weeks to induce liver fibrosis. **j**, Fold change in serum levels of suPAR. Representative results of two independent experiments (vehicle, $n = 4$; CCl_4, $n = 9$ mice per group). Two-tailed unpaired Student's *t*-test. **k**, Co-immunofluorescence staining and quantification of uPAR (red) and Ki-67 (green). Representative results of two independent experiments ($n = 2$ mice per group). **l**, Co-immunofluorescence staining and quantification of uPAR (red) and IL-6 (green). Representative results of two independent experiments ($n = 3$ mice per group). **m**, Representative SA-β-gal staining. Representative results of two independent experiments ($n = 3$ mice per group). Data are mean ± s.e.m. (**c**, **d**, **h**, **j**, **l**).

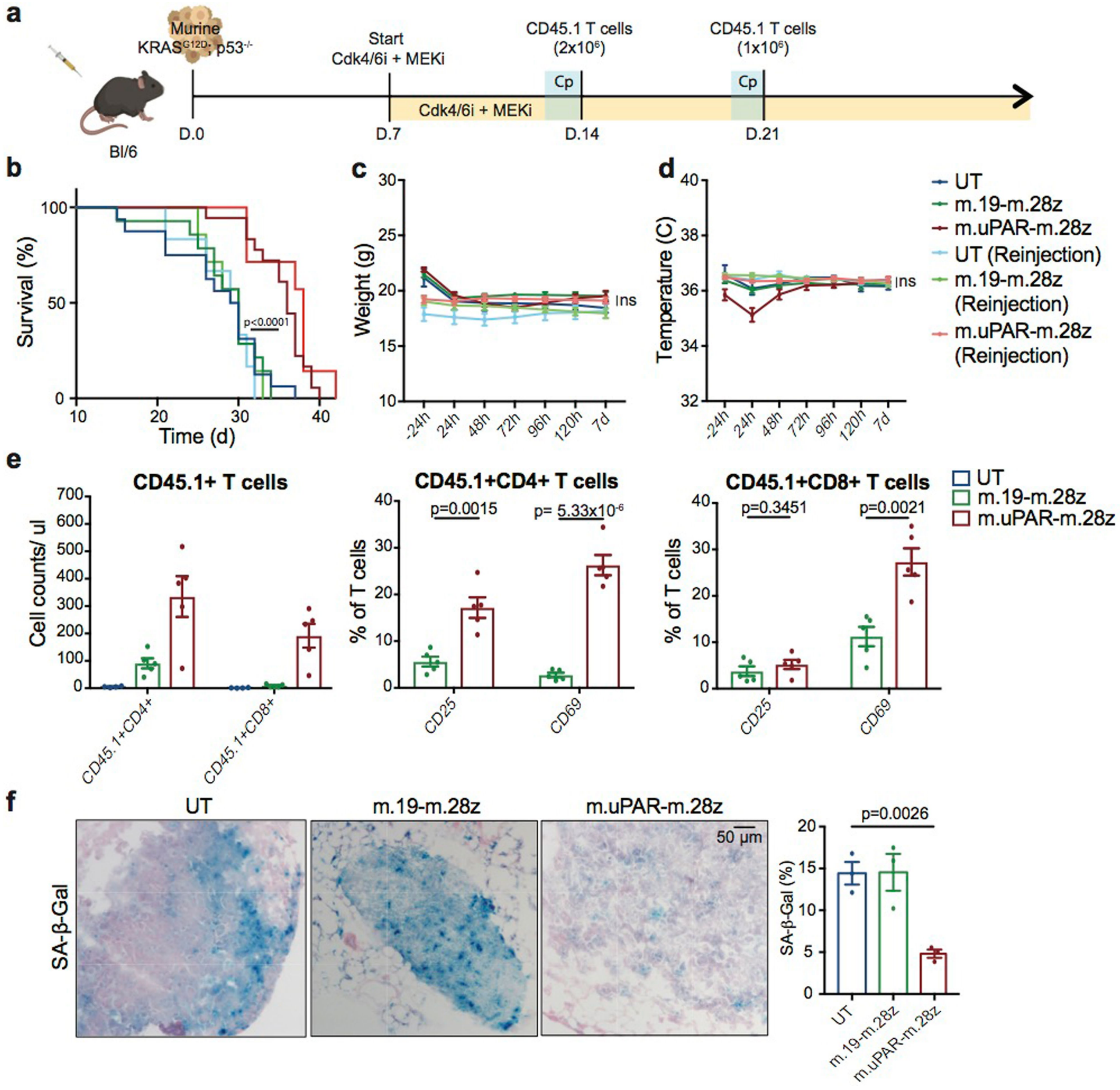


Extended Data Fig. 3 | uPAR is a marker of senescence in senescence-associated human pathologies. **a**, Left, immunohistochemical expression of human uPAR (h.uPAR) and SA-β-gal in human samples of hepatitis-induced liver fibrosis (*n* = 7 patients). Right, co-immunofluorescence staining and quantification of uPAR (red) and p16 (green) or uPAR (red) and IL-6 (green) in human samples of hepatitis-induced liver fibrosis (*n* = 3). **b**, Left, immunohistochemical expression of uPAR and SA-β-gal in human samples from patients with eradicated hepatitis C virus (HCV) and residual liver fibrosis (*n* = 7 patients). Right, co-immunofluorescence staining and quantification of uPAR (red) and p16 (green) or uPAR (red) and IL-6 (green) in human samples of HCV-induced liver fibrosis (*n* = 3). Data are mean ± s.e.m. (**a**, **b**). **c**, Immunohistochemical staining of uPAR in human carotid endarterectomy samples (*n* = 5 patients). **d**, Immunohistochemical staining of uPAR in human pancreas bearing pancreatic intraepithelial neoplasia (PanIN) compared to normal pancreas controls (*n* = 3 patients).

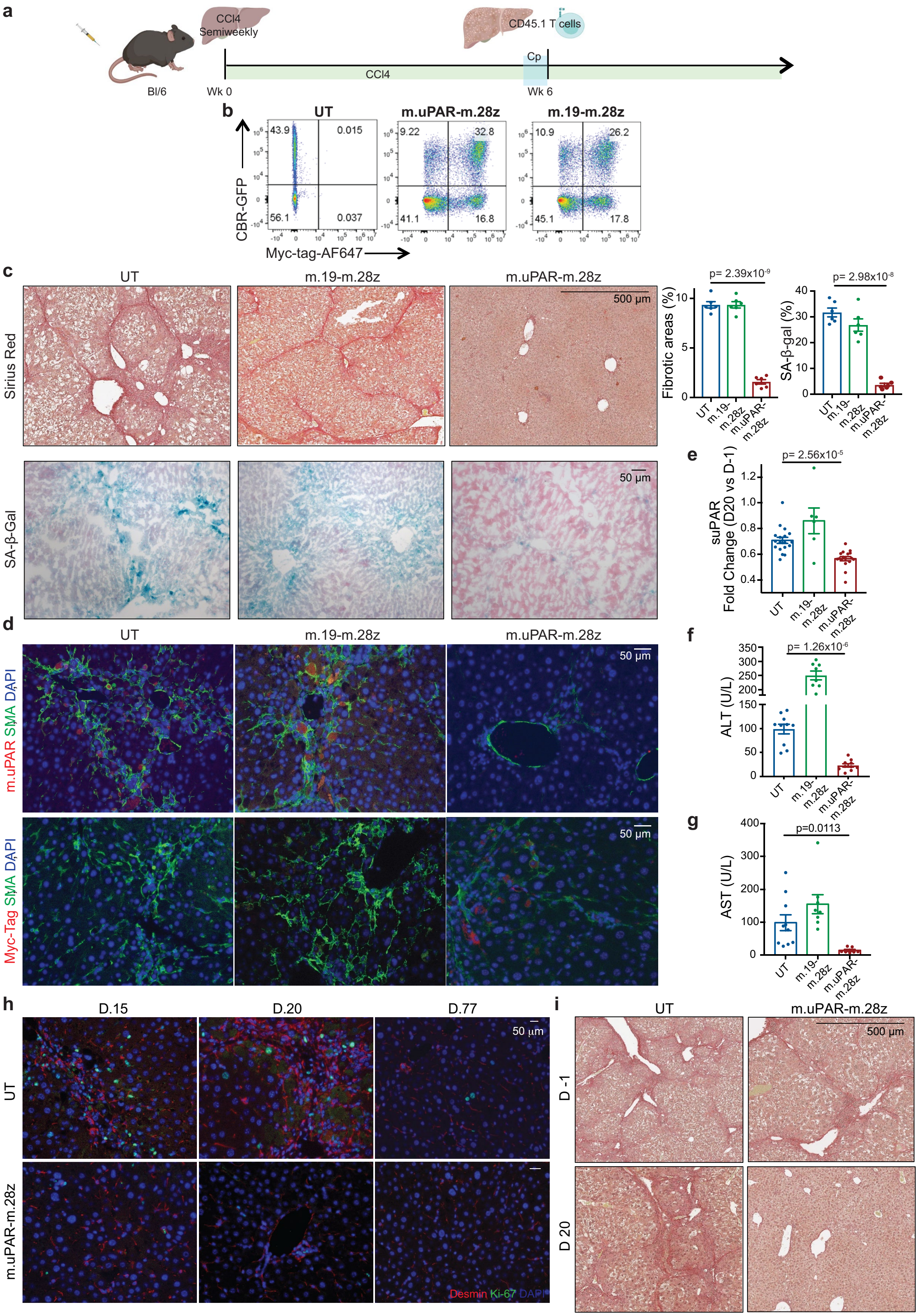


Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | m.uPAR-h.28z CAR T cells selectively target uPAR-positive cells. a, Construct maps encoding human m.uPAR-h.28z and h.19-h.28z CARs or mouse m.uPAR-m.28z and m.19-m.28z CARs. **b**, Flow cytometry analysis showing the expression levels of CAR and LNGFR in m.uPAR-h.28z and h.19-h.28z CAR T cells compared to untransduced T cells. Representative results of *n* = 4 independent experiments. **c**, Flow cytometry analysis of mouse uPAR and human CD19 expression on wild-type NALM6 cells and NALM6-m.uPAR cells. Representative results of *n* = 3 independent experiments. **d**, Cytotoxic activity of m.uPAR-h.28z, h.19-h.28z and untransduced T cells as determined by 4-h calcein assay with firefly luciferase (FFL)-expressing NALM6 wild-type or NALM6-m.uPAR cells as targets. Representative results of *n* = 3 independent experiments performed in triplicate. Data are mean ± s.e.m. **e**, Granzyme B (GrB) and IFNγ expression of CD4$^+$ and CD8$^+$ m.uPAR-h.28z CAR T cells 18 h after co-culture with wild-type NALM6, NALM6-m.uPAR or senescent KP cells as determined by intracellular cytokine staining. Results of *n* = 1 independent experiment (no target and NALM6 WT, *n* = 2; NALM6-m.uPAR and KP senescent, *n* = 3 replicates). Data are mean ± s.e.m. **f**, Experimental layout for Fig. 2c–i. Mice were injected with a plasmid encoding Nras^{G12V}-GFP-luciferase and treated with 0.5×10^6 m.uPAR-h.28z CAR T cells or untransduced T cells 10 days after injection. Mice were euthanized 15 days after CAR administration and livers were used for further analysis. Images were created with BioRender.com. **g**, Flow cytometry analysis of mouse uPAR and human CD19 expression on wild-type Eμ-ALL01 cells and Eμ-ALL01-m.uPAR cells. Representative results of *n* = 3 independent experiments. **h**, Flow cytometry staining of Myc-tag and mouse uPAR on m.uPAR-m.28z CAR T cells, m.19-m.28z CAR T cells and untransduced T cells as compared to FMO control. Representative results of *n* = 2 independent experiments.

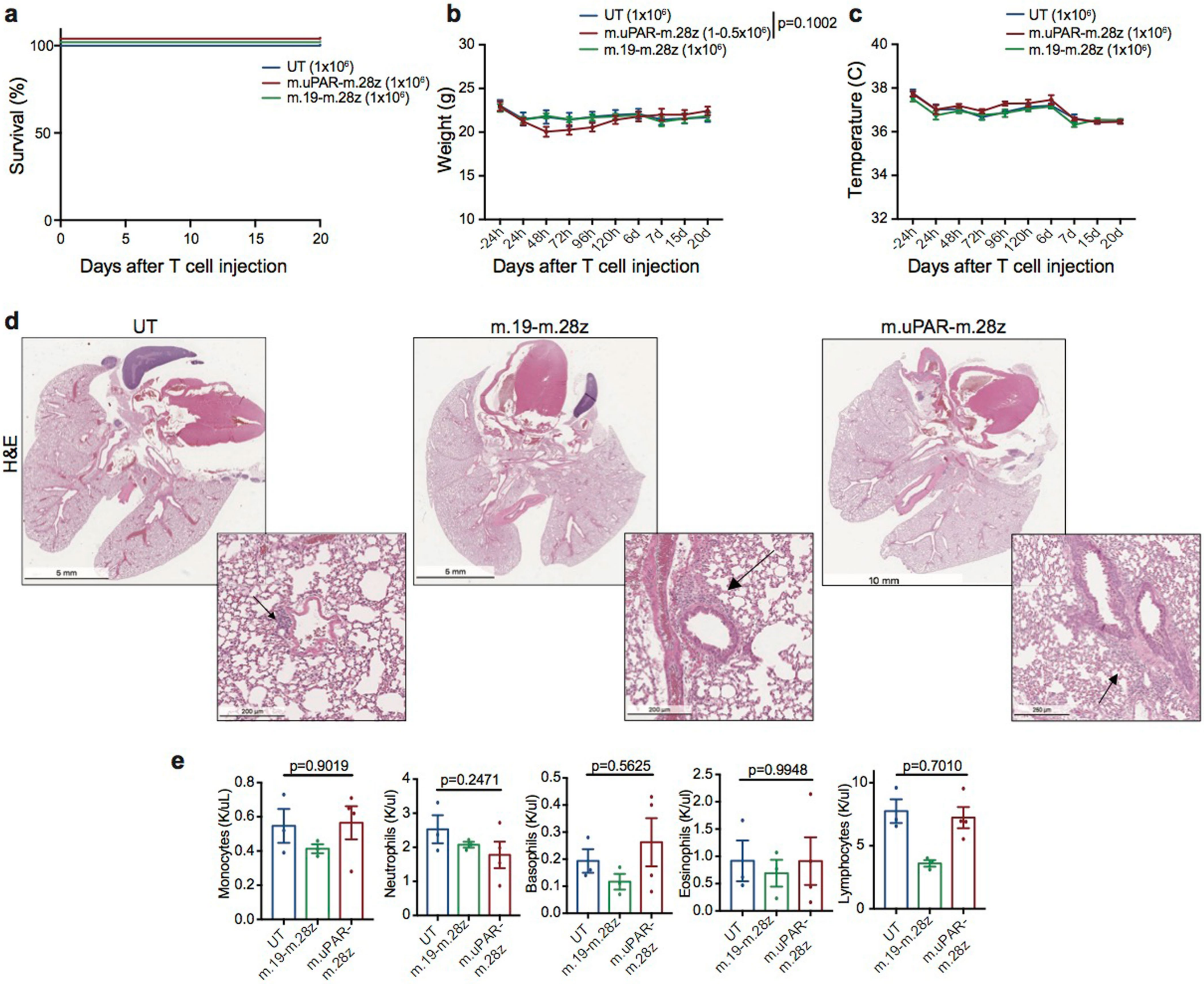


Extended Data Fig. 5 | Senolytic CAR T cells target senescent cells in a *Kras*^{G12D}-driven model of lung cancer. a, Experimental layout. C57BL/6N mice were intravenously injected with 10,000 *Kras*^{G12D};*p53*$^{-/-}$ cells. Treatment with combined MEK inhibitor (1 mg per kg body weight) and CDK4/6 inhibitors (100 mg per kg body weight) was started seven days later, followed by adoptive transfer of 2 × 10^6 CD45.1$^+$ T cells (m.uPAR-m.28z CAR T cells, m.19-m.28z CAR T cells or untransduced T cells) one week later. A subset of mice received a second infusion of 1 × 10^6 m.uPAR-m.28z CAR T cells, m.19-m.28z CAR T cells or untransduced T cells seven days after the first injection of T cells. The images of the mouse, tumour cells and CAR T cells were created with BioRender.com. Cp, cyclophosphamide. **b**, Kaplan–Meier curve showing survival of mice (one-sided log-rank (Mantel–Cox) test). Results of two independent experiments (UT, *n* = 16; m.19-m.28z, *n* = 14; m.uPAR-m.28z, *n* = 18; UT reinjection, *n* = 6; m.19-m.28z reinjection, *n* = 7; m.uPAR-m.28z reinjection, *n* = 7 mice). **c**, **d**, Weight (**c**) and temperature (**d**) measured 24 h before and at different time points after CAR T cell infusion. *P* values (ns, not significant) refer to the comparison between untransduced and m.uPAR-m.28z injected mice at 48 h (weight, *P* = 0.9329; temperature, *P* = 0.1534). Results of one independent experiment (UT, *n* = 5; m.19-m.28z, *n* = 5; m.uPAR-m.28z, *n* = 8; UT reinjection, *n* = 5; m.19-m.28z reinjection, *n* = 7; m.uPAR-m.28z reinjection, *n* = 7 mice). **e**, Cell counts of CD45.1$^+$ T cells and expression of the activation markers CD25 and CD69 (UT, *n* = 4; m.19-m.28z, *n* = 5; m.uPAR-m.28z, *n* = 5 mice) on CD45.1$^+$ T cells in the lungs of mice seven days after administration of m.uPAR-m.28z CAR T cells, m.19-m.28z CAR T cells or untransduced T cells. **f**, Representative SA-β-gal staining and quantification in the lungs of mice seven days after treatment with m.uPAR-m.28z CAR T cells compared to mice that were treated with m.19-m.28z CAR T cells or untransduced T cells (*n* = 3 mice per group). Data are mean ± s.e.m.; two-tailed unpaired Student's *t*-test. (**c**–**f**).

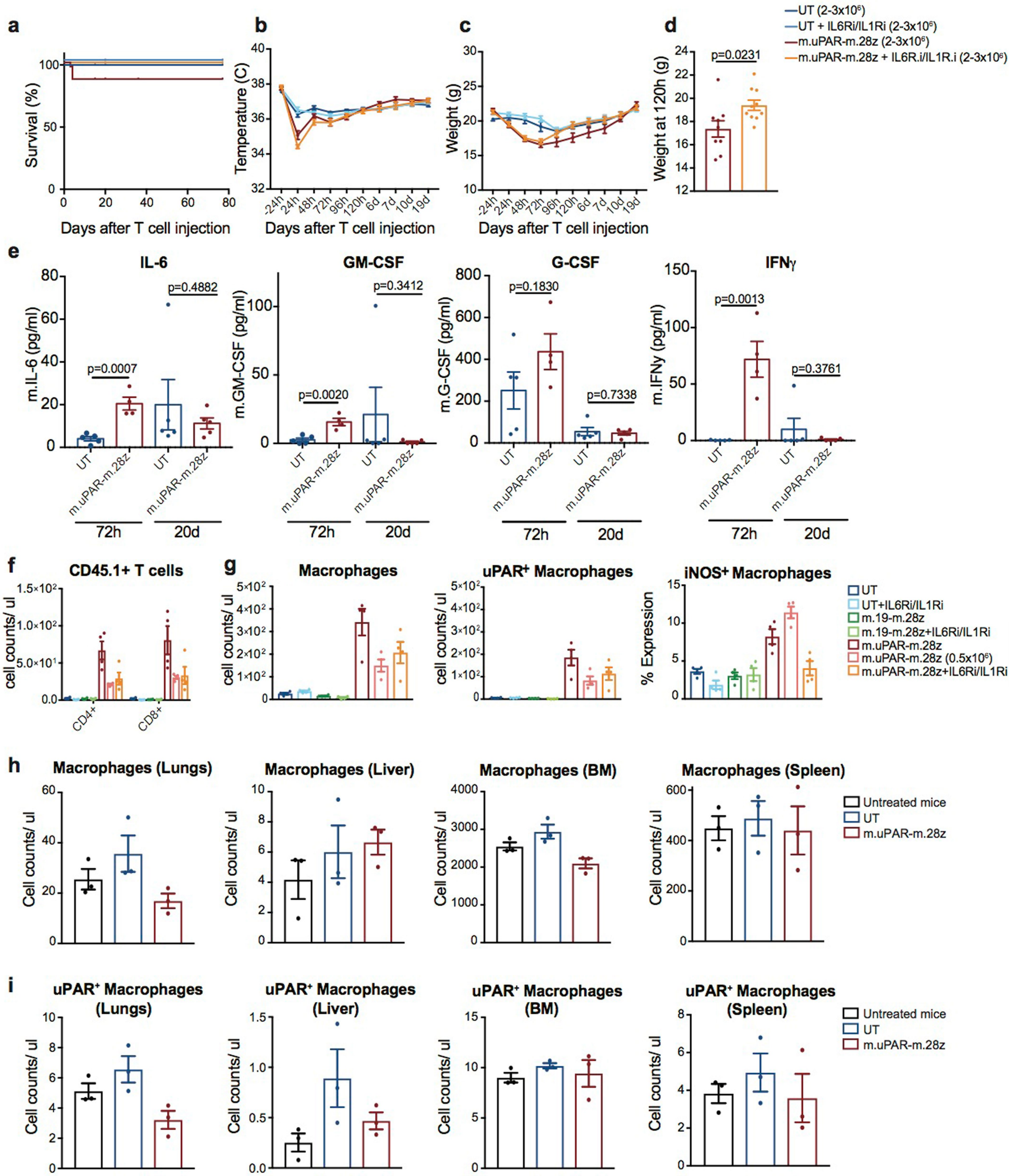


Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Senolytic CAR T cells show therapeutic activity in CCl_4-induced liver fibrosis. a, Layout for experiments performed using the CCl_4-induced liver fibrosis model: C57BL/6N mice received intraperitoneal infusions of CCl_4 twice weekly for six weeks and were intravenously infused with $0.5–1 \times 10^6$ (Fig. 3) or $2–3 \times 10^6$ (**c**–**i**) mouse m.uPAR-m.28z CAR T cells, m.19-m.28z CAR T cells or untransduced T cells 16–24 h after administration of cyclophosphamide (200 mg kg^{-1}). Mice were euthanized 20 days after CAR T cell infusion to assess liver fibrosis. Images were created with BioRender.com. **b**, Expression of GFP-tagged click beetle red (CBR) luciferase and Myc-tag in m.uPAR-m.28z and m.19-m.28z CAR T cells that were used for T cell imaging experiments (Fig. 3g, h) compared to control T cells. Representative results of *n* = 2 independent experiments. **c**, Sirius red and SA-β-gal staining and quantifications in livers from treated mice (*n* = 6 mice per group). **d**, Co-immunofluorescence of uPAR (red) and SMA (green) or Myc-tag (red) and SMA (green) in the livers of treated mice. **e**, Fold change difference in serum levels of suPAR 20 days after compared to 1 day before (day −1) injection of CAR T cells (UT, *n* = 18; m.19-m.28z, *n* = 6; m.uPAR-m.28z, *n* = 17 mice). **f**, **g**, Levels of serum ALT (**f**) and AST (**g**) 20 days after CAR treatment (UT, *n* = 10; m.19-m.28z, *n* = 8; m.uPAR-m.28z, *n* = 10 mice). **h**, Co-immunofluorescence staining of desmin (red) and Ki-67 (green) in the livers of mice 15, 20 and 77 days after treatment with CAR T cells. CCl_4 treatment was stopped 20 days after T cell infusion (*n* = 3 mice per group). **i**, Mice were treated with CCl_4 for 10 weeks. Sirius red staining in the livers of mice before (day −1) and 20 days after T cell administration (UT, *n* = 4; m.uPAR-m.28z, *n* = 2 mice). Representative results of *n* = 2 independent experiments (**c**–**i**). Data are mean ± s.e.m.; two-tailed unpaired Student's *t*-test (**c**, **e**–**g**).

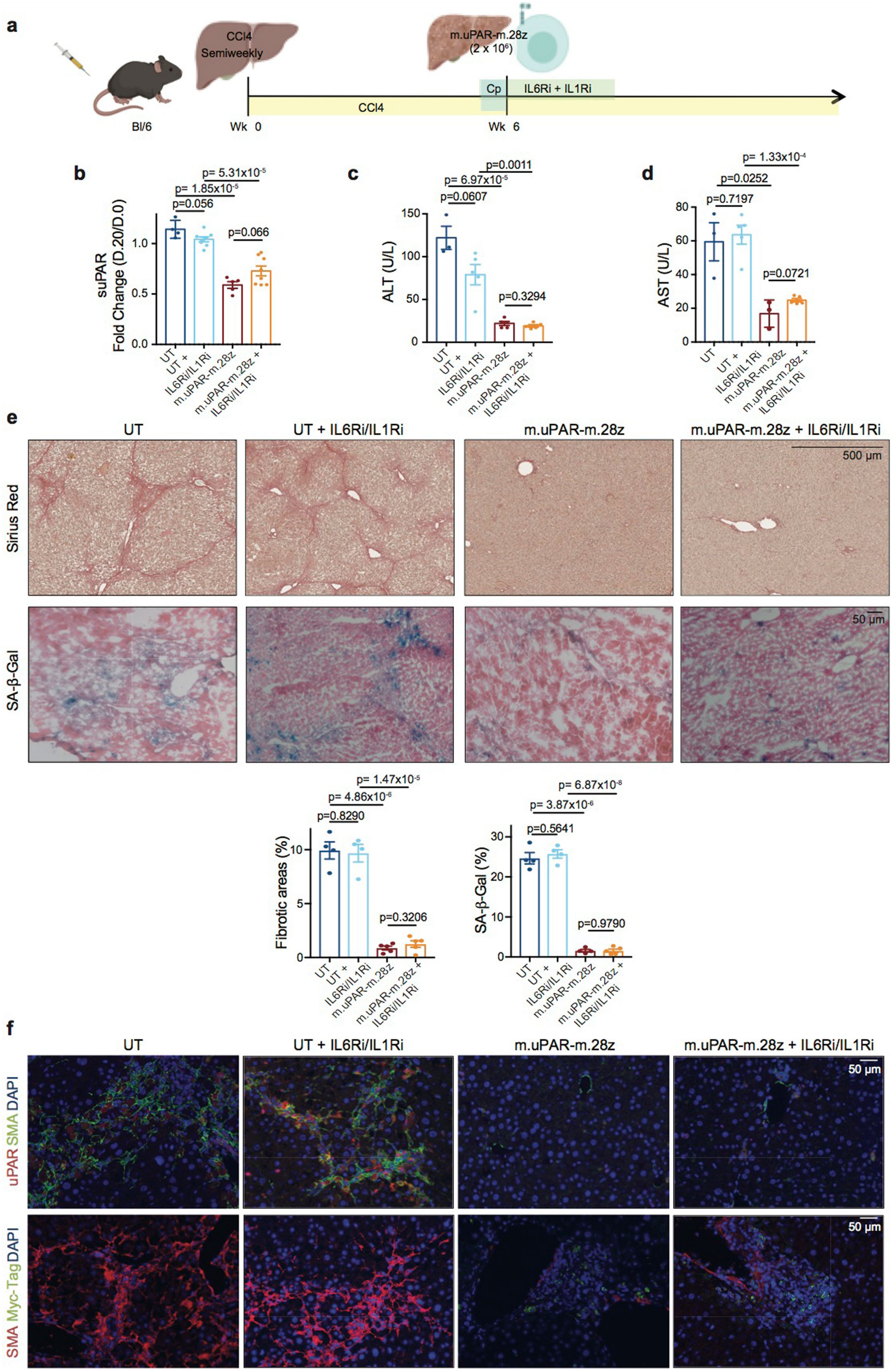


Extended Data Fig. 7 | Safety profile of m.uPAR-m.28z CAR T cells at therapeutic doses of T cells. **a**–**e**, C57BL/6N mice received intraperitoneal infusions of CCl_4 twice weekly for six weeks and were intravenously injected with $0.5–1 \times 10^6$ m.uPAR-m.28z CAR T cells, 1×10^6 m.19-m.28z CAR T cells or untransduced T cells 16 h after administration of cyclophosphamide (200 mg kg^{-1}). Mice were euthanized 20 days after T cell administration to assess potential toxicities and lung histopathology. **a**, Kaplan–Meier curve showing survival of mice after treatment with m.uPAR-m.28z CAR T cells (*n* = 16 mice), m.19-m.28z CAR T cells (*n* = 6 mice) or untransduced T cells (*n* = 6 mice). **b**, **c**, Weight (**b**) and temperature (**c**) of mice measured before and at different time points after CAR T cell infusion (UT and m.19-m.28z, *n* = 6; m.uPAR-m.28z, *n* = 7 mice). The *P* value in **b** refers to differences in weight at 48 h. **d**, **e**, Representative H&E staining of lungs (**d**) and complete blood counts (**e**) of treated mice 20 days after T cell infusion (UT and m.19-m.28z, *n* = 3 or 4; m.uPAR-m.28z, *n* = 4 mice). An increased accumulation of macrophages was observed in the immunodeficient setting. Representative results of *n* = 1 independent experiment (**a**–**e**). Data are mean ± s.e.m. (**b**, **c**, **e**); two-tailed unpaired Student's *t*-test (**b**, **e**).



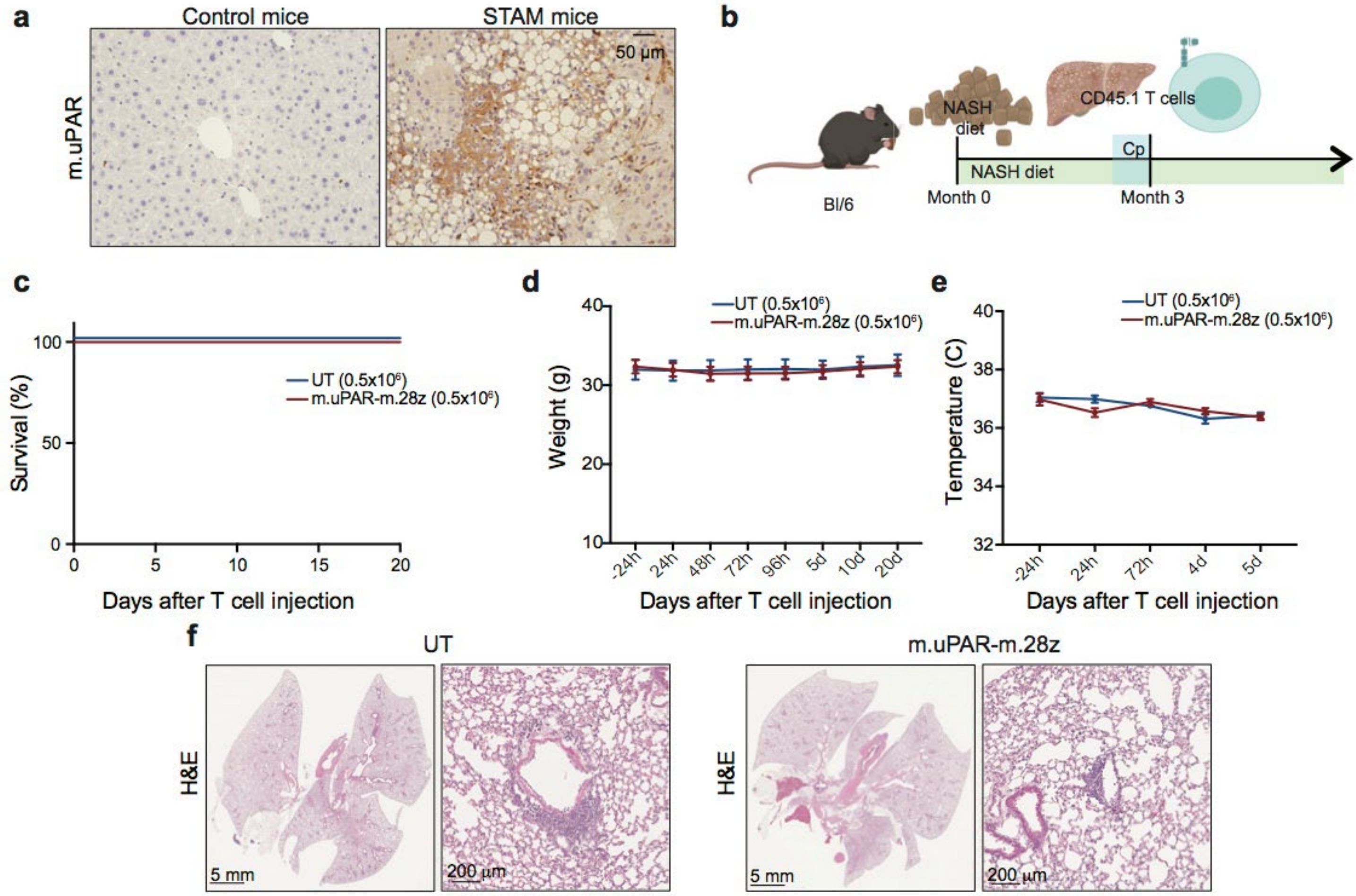
Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Safety profile of m.uPAR-m.28z CAR T cells at supratherapeutic doses of T cells. C57BL/6N mice received intraperitoneal infusions of CCl_4 twice weekly for six weeks followed by intravenous infusion of 2–3 × 10^6 m.uPAR-m.28z CAR T cells or untransduced T cells 16–24 h after administration of cyclophosphamide (200 mg kg^{-1}). A subset of mice (as specified in the figure) received additional treatment with IL-6R-blocking antibodies (IL6Ri) and the IL-1R antagonist anakinra (IL1Ri), starting 24 h before T cell infusion and continuing daily until 6 days after T cell infusion. Mice were euthanized 12 weeks after CAR infusion to assess potential toxicities. **a**, Kaplan–Meier curve showing survival of mice after injection of CAR T cells (UT, $n = 19$; UT + IL6Ri/IL1Ri, $n = 7$; m.uPAR-m.28z, $n = 30$; m.uPAR-m.28z + IL6Ri/IL1Ri, $n = 19$ mice). **b**, **c**, Temperature (**b**) and weight (**c**) of treated mice (UT, $n = 7$; UT + IL6Ri/IL1Ri, $n = 8$; m.uPAR-m.28z, $n = 11$; m.uPAR-m.28z + IL6Ri/IL1Ri, $n = 10$ mice). **d**, Weight of mice 120 h after infusion with either m.uPAR-m.28z or m.uPAR-m.28z CAR T cells and additional treatment with IL6Ri and IL1Ri (m.uPAR-m.28z, $n = 11$; m.uPAR-m.28z + IL6Ri/IL1Ri, $n = 10$ mice). **e**, Serum levels of IL-6, GM-CSF, G-CSF and IFNγ in mice that were treated with either m.uPAR-m.28z or untransduced T cells 72 h or 20 days after T cell infusion (UT, $n = 5$; m.uPAR-m.28z, $n = 4$ mice at 72 h; $n = 5$ mice at 20 days). **f**, **g**, Number of adoptively transferred CD45.1$^+$ T cells (**f**) and number of macrophages, uPAR$^+$ and iNOS$^+$ macrophages (**g**) in the lungs of mice that were treated with m.uPAR-m.28z CAR T cells, m.19-m.28z CAR T cells or untransduced T cells alone or in combination with treatment with IL6Ri and ILR1Ri three days after T cell infusion ($n = 4$ mice per group). **h**, **i**, Number of macrophages (**h**) and uPAR$^+$ macrophages (**i**) in the lungs, liver, bone marrow (BM) and spleen of untreated mice or mice treated with either m.uPAR-m.28z CAR T cells or untransduced T cells 12 weeks after T cell infusion ($n = 3$ mice per group). Representative results of $n = 3$ independent experiments (**a**–**d**) or $n = 1$ independent experiment (**e**–**i**). All data are mean ± s.e.m.; two-tailed unpaired Student's *t*-test (**d**, **e**).

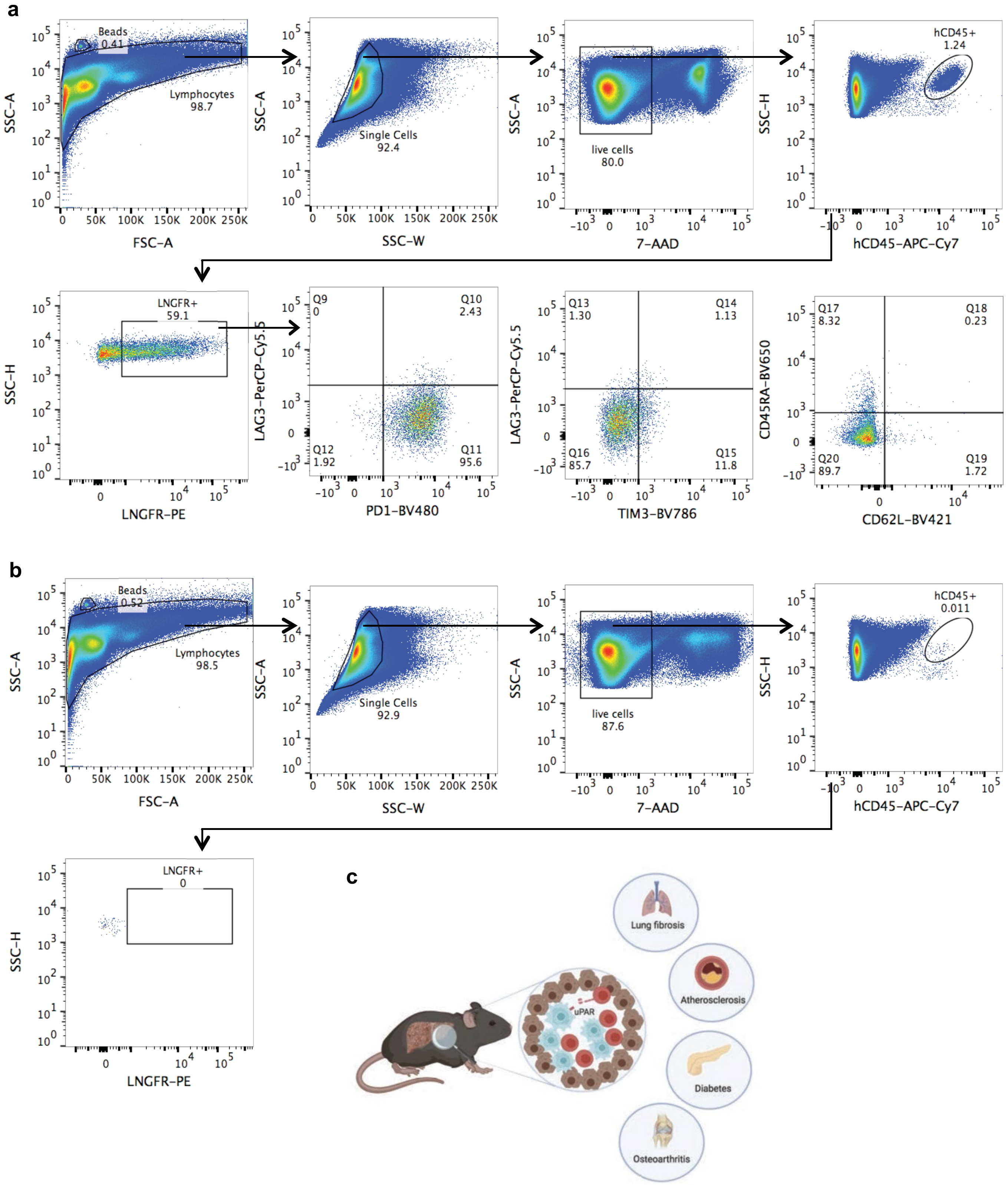


Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Therapeutic intervention with IL-6R and IL-1R inhibitors does not decrease the therapeutic efficacy of senolytic CAR T cells in CCl_4-induced liver fibrosis. a, Experimental layout. C57BL/6N mice received intraperitoneal infusions of CCl_4 twice weekly for six weeks and were intravenously infused with $2–3 \times 10^6$ m.uPAR-m.28z CAR T cells or untransduced T cells 24 h after administration of cyclophosphamide (200 mg kg^{-1}). IL-6R-blocking antibodies (IL6Ri) and anakinra (ILRi) were first administered 24 h before T cell infusion followed by daily (IL6Ri) or twice daily (IL1Ri) injections for the first six days until treatment was stopped. Mice were euthanized 20 days after T cell infusion to assess liver fibrosis. Images were created with BioRender.com. **b**, Fold change difference in serum levels of suPAR 20 days after compared to 1 day before (day −1) CAR T cell treatment (UT, *n* = 4; UT + IL6Ri/IL1Ri, *n* = 8; m.uPAR, *n* = 5; m.uPAR + IL6Ri/IL1Ri, *n* = 8 mice). **c**, **d**, Levels of serum ALT (**c**) and AST (**d**) in treated mice 20 days after T cell infusion (UT, *n* = 3; UT + IL6Ri/IL1Ri, *n* = 5; m.uPAR-m.28z, *n* = 5 (ALT) and *n* = 3 (AST); m.uPAR-m.28z + IL6Ri/IL1Ri, *n* = 5 mice). **e**, Representative levels of fibrosis evaluated by Sirius red staining and SA-β-gal staining in livers from treated mice and quantification of liver fibrosis and SA-β-gal^+ cells in the respective livers 20 days after treatment (UT, *n* = 4; UT + IL6Ri/IL1Ri, *n* = 4; m.uPAR-m.28z, *n* = 4; m.uPAR-m.28z + IL6Ri/IL1Ri, *n* = 5 mice). **f**, Co-immunofluorescence staining of uPAR (red) and SMA (green) or Myc-tag (red) and SMA (green) in the livers of treated mice. Representative results of *n* = 1 independent experiment (**b**–**f**). Data are mean ± s.e.m.; two-tailed unpaired Student's *t*-test (**b**–**e**).



Extended Data Fig. 10 | Safety profile of senolytic CAR T cells at therapeutic doses in a mouse model of NASH-induced liver fibrosis. **a**, Immunohistochemical expression of uPAR in samples from the 'STAM' model[52,58] ($n = 3$ mice). **b**, Experimental layout for experiments performed using the model of diet-induced NASH (Fig. 4, this figure). C57BL/6N mice were treated with a chow or a NASH-inducing diet[50] for three months, followed by intravenous infusion with 0.5×10^6 m.uPAR-m.28z CAR T cells or untransduced T cells 16 h after administration of cyclophosphamide (200 mg kg^{-1}). Mice were euthanized 20 days after CAR infusion to assess liver fibrosis. Images were created with BioRender.com. **c**, Kaplan–Meier curve showing survival of mice after treatment with either m.uPAR-m.28z CAR T cells or untransduced T cells (m.uPAR-m.28z, $n = 16$; UT, $n = 10$ mice). **d**, **e**, Weight (**d**) and temperature (**e**) of mice 24 h before and at different time points after T cell infusion (m.uPAR-m.28z, $n = 11$; UT, $n = 9$ mice). Data are mean ± s.e.m. **f**, Representative H&E staining of the lungs of treated mice (m.uPAR-m.28z, $n = 6$; UT, $n = 4$ mice). Representative results of $n = 2$ independent experiments (**c**–**f**).



Extended Data Fig. 11 | Gating strategies, summary and potential applications of senolytic CAR T cells. a, **b**, Representative flow cytometry staining of m.uPAR-h.28z CAR T cells (**a**) or untransduced T cells (**b**) obtained from the livers of mice that had undergone HTVI (as depicted in Fig. 2). Representative results of one independent experiment (*n* = 4 mice per group). **c**, Summary of the key points of our findings. uPAR-28z CAR T cells (red) infiltrate fibrotic livers that contain senescent cells (blue) and efficiently eliminate them, leading to fibrosis resolution and improved liver function. The therapeutic action of senolytic uPAR-28z CAR T cells might be extended to other senescence-associated diseases such as atherosclerosis, diabetes or osteoarthritis. Images were created with BioRender.com.

nature research

Corresponding author(s): Scott W Lowe
Last updated by author(s): April 22, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	BD LSR-II and BD LSR-Fortessa cytometer, Cytek Aurora (CYTEK), Xenogen IVIS Imaging System living image V4.4, Microsoft Excel for Mac 2011.
Data analysis	FlowJo 10.1 , GraphPad Prism V6 and V7, Living Image 4.4, Image J, version 2.0.0-rc-43/1.51h. RNAseq analysis was performed with the following software: HTSeq v0.5.3, picard tools v1.124, R v3.2.0, STAR v2.5.0a, samtools v0.1.19, Microsoft Excel for Mac 2011.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The RNA-seq data has been deposited in the Gene Expression Omnibus (GEO) under the accession number GSE145642.
The datasets generated during the current study are available from the corresponding author upon reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to pre-determine sample size. Sample sizes were estimated based on preliminary experiments, with an effort to achieve a minimum of n=5 mice per treatment group which proved to be sufficient to reproducibly observe a statistically significant difference.
Data exclusions	No data were excluded throughout the studies.
Replication	All in vitro and in vivo experiments were repeated in replicates and/or from different subjects in independent experiments. All attempts at replication were successful. Efficacy of CAR T cell treatment may vary between donors.
Randomization	Senescence burden for HTVI was determined by bioluminescent imaging one day prior to CAR T cell transfer and by suPAR measurement in the liver fibrosis model. Since senescent burdens were very even, mice were randomly assigned into treatment groups. Buffy coats were obtained from anonymous donors.
Blinding	Mouse conditions were observed by an operator who was blinded to the treatment groups in addition to the main investigator who was not blind to group allocation. Analysis of data was not performed in blinded fashion. Data analysis are based on objectively measurable data (fluorescence intensity, cell count, blood tests).

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	The following fluorophore-conjugated antibodies were used ("h" prefix denotes anti-human, "m" prefix denotes anti-mouse): hCD45 APC-Cy7 (clone 2D1, BD, #557833, Lot: 9081815, 1:100), hCD4 BUV395 (clone SK3, BD, #563550, Lot: 6252529, 1:100), hCD4 BV480 (clone SK3, BD, #566104, Lot: 8092993, 1:50.), hCD62L BV421 (clone DREG-56, BD, #563862, Lot: 8194954, 1:100.), hCD45RA BV650 (clone HI100, BD, #563963, Lot: 9057952, 1:100.), hPD1 BV480 (clone EH12.1, BD, #566112, Lot: 8235507, 1:100.), hCD19 BUV737 (clone SJ25C1, BD, #564303, Lot: 8130572, 1:100.), hCD271 PE (clone C40-1457, BD, #557196, Lot: 7068641, 1:100.), hIL2 PE-Cy7 (clone MQ1-17H12, Invitrogen, #25-7029-42, Lot: 4336863, 1:50.), hTNFa BV650 (clone Mab11, BD, #563418, Lot: 7082880, 1:50.), hIFNg BUV395 (clone B27, BD, #563563, Lot: 6320836, 1:50.) hTIM3 BV785 (clone F38-2E2, Biolegend, #345032, Lot: B265346, 1:100.), hCD8 PE-Cy7 (clone SK1, eBioscience, #25-0087-42, Lot: 2066348, 1:100.), hCD8 APC-Cy7 (clone SK1, BD, #557834, Lot: 7110951, 1:50.), hCD223 PerCP-eFluor710 (clone 3DS223H, eBioscience, #46-2239-42, Lot: 4321735, 1:100.), hGrB APC (clone GB12, Invitrogen, #MHGB05, Lot: 1884625, 1:67.), hMyc-tag AF647 (clone 9B11, Cell Signaling Technology, #2233S, Lot: 23, 1:50.), hCD19 PB (clone SJ25-C1, Invitrogen, #MHCD1928, 1:100.), hCD87 APC (clone VIM5, eBioscience, #17-3879-42, Lot: 17-3879-42, 1:50.), hCD87 PerCp-eFluor710 (clone VIM5, eBioscience, #46-3879-42, Lot: 46-2239-42, 1:50.), muPAR PE (R&D Systems, FAB531P, Lot: ABLH0419081, 1:50.), muPAR AF700 (R&D Systems, FAB531N, Lot: 1552229, 1:50.), mCD45.1 APC-Cy7 (clone A20, Biolegend, #110716, Lot: B285685, 1:200.), m.CD45.1 BV785 (clone A20, Biolegend, #110743, Lot: B270183, 1:100.), mCD45.2 PE (clone 104, Biolegend, #109808, Lot: B271929, 1:100.), mCD45.2 AF700 (clone 104, Biolegend, #109822, Lot: B252126, 1:200.), mSiglec-F PerCP-Cy5.5 (clone E50-2440, BD, #565526, Lot: 8232650, 1:200.), mI-A/I-E BV605 (clone M5/114.15.2, Biolegend, #107639, Lot: B293222, 1:50.), mF4/80 BV421 (clone: T45-2342, BD, #565411, Lot: 8330526, 1:200.), mCD11b BUV395 (clone: M1/70, BD, #563553, Lot: 8339988, 1:200.), mCD11c BV650 (clone: N418, Biolegend, #117339, Lot: B253523, 1:200.), mLY6G BV510 (clone: 1A8, Biolegend, #127633, Lot: B266675, 1:200.), mLY6G

APC/Fire750 (clone: 1A8, Biolegend, #127652, Lot: B274284, 1:100.), miNOS PE-Cy7 (clone: CXNFT, eBioscience, #25-5920-82, Lot: 2127491, 1:200.), mCD19 PE (clone 1D3/CD19, Biolegend, #152408, Lot: B260181, 1:100.), mCD25 BV605 (clone: PC61, Biolegend, #102035, Lot: B291215, 1:50.), mCD69 PerCpCy5.5 (clone: H1.2F3, Biolegend, #104522, Lot: B244018, 1:100.), mCD3 AF488 (clone: 17A2, Biolegend, #100210, Lot: B284975, 1:100.), mCD4 BUV395 (clone: GK1.5, BD, #563790, Lot: 9101822, 1:50.), mCD4 FITC (clone: GK1.5, BD, #553729, Lot: 9204449, 1:50.), mCD8 PE-Cy7 (Clone: 53-6.7, Biolegend, #100722, Lot: B282418, 1:50.),Alexa Fluor 647 AffiniPure F(ab)2 Fragment Goat Anti-Rat IgG (Jackson ImmunoResearch, #112-6606-072), huPAR (R&D. AF807 Lot.BBS0318071. 1:50), muPAR (R&D. AF534 Lot.DCL0418021. 1:50), mNRAS (Santa Cruz. SC-31 Lot.A1020. 1:50), mSMA (abcam. Ab5694 Lot.GR283004-16. 1:50), mKATE (Evrogen. ab233 Lot.23301201267. 1:1000), hCD3 (abcam. ab5690 Lot.GR3220039-4. 1:100), myc-tag (Cell Signaling. 2276S Lot.24. 1:50), mKi-67 (abcam, ab16667 Lot.GR3305281-1. 1:200), mIL-6 (abcam. ab6672 Lot.GR3195128-19. 1:50), p16-INK4A (Proteintech. 10883-1-AP Lot.00057396. 1:50), mP-ERKT202/Y204 (Cell Signaling.4370 Lot.15. 1:800), desmin (ThermoFisher Scientific . RB-9014 Lot.9014p1806Q. 1:200), AF488 donkey anti-rabbbit (Invitrogen. A21206 Lot.1874771. 1:500), AF488 donkey anti-mouse (Invitrogen. A21202 Lot.1820538. 1:500), AF594 donkey anti-rabbit (Invitrogen. A21207 Lot.1602780. 1:500), AF594 donkey anti-mouse (Invitrogen A21203 Lot.1163390. 1:500), AF594 donkey anti-goat (Invitrogen. A11058 Lot.2045324. 1:500), AF594 goat anti-rat (Invitrogen A11007 Lot.1903506. 1:500).

Validation

All used antibodies were titrated. All the antibodies are validated for use in flow cytometry or immunohistochemistry or immunofluorescence. Data are available on the manufacturer's website. The following primary antibodies have been validated by the manufacturer in the mentioned species: hCD45 APC-Cy7 (clone 2D1, BD, #557833, Human), hCD4 BUV395 (clone SK3, BD, #563550, Human), hCD4 BV480 (clone SK3, BD, #566104, Human), hCD62L BV421 (clone DREG-56, BD, #563862, Human), hCD45RA BV650 (clone HI100, BD, #563963, Human), hPD1 BV480 (clone EH12.1, BD, #566112, Human), hCD19 BUV737 (clone SJ25C1, BD, #564303, Human), hCD271 PE (clone C40-1457, BD, #557196, Human), hIL2 PE-Cy7 (clone MQ1-17H12, Invitrogen, #25-7029-42, Human), hTNFa BV650 (clone Mab11, BD, #563418, Human), hIFNg BUV395 (clone B27, BD, #563563, Human), hTIM3 BV785 (clone F38-2E2, Biolegend, #345032, Human), hCD8 PE-Cy7 (clone SK1, eBioscience, #25-0087-42, Human), hCD8 APC-Cy7 (clone SK1, BD, #557834, Human), hCD223 PerCP-eFluor710 (clone 3DS223H, eBioscience, #46-2239-42, Human), hGrB APC (clone GB12, Invitrogen, #MHGB05, Human), hMyc-tag AF647 (clone 9B11, Cell Signaling Technology, #2233S, Transfected mammalian cells), hCD19 PB (clone SJ25-C1, Invitrogen, #MHCD1928, Human), hCD87 APC (clone VIM5, eBioscience, #17-3879-42, Human), hCD87 PerCp-eFluor710 (clone VIM5, eBioscience, #46-3879-42, Human), muPAR PE (R&D Systems, FAB531P, Mouse), muPAR AF700 (R&D Systems, FAB531N, Mouse), mCD45.1 APC-Cy7 (clone A20, Biolegend, #110716, Mouse), m.CD45.1 BV785 (clone A20, Biolegend, #110743, Mouse), mCD45.2 PE (clone 104, Biolegend, #109808, Mouse), mCD45.2 AF700 (clone 104, Biolegend, #109822, Mouse), mSiglec-F PerCP-Cy5.5 (clone E50-2440, BD, #565526, Mouse), mI-A/I-E BV605 (clone M5/114.15.2, Biolegend, #107639, Mouse), mF4/80 BV421 (clone: T45-2342, BD, #565411, Mouse), mCD11b BUV395 (clone: M1/70, BD, #563553, Mouse), mCD11c BV650 (clone: N418, Biolegend, #117339, Mouse), mLY6G BV510 (clone: 1A8, Biolegend, #127633, Mouse), mLY6G APC/Fire750 (clone: 1A8, Biolegend, #127652, Mouse), miNOS PE-Cy7 (clone: CXNFT, eBioscience, #25-5920-82, Mouse), mCD19 PE (clone 1D3/CD19, Biolegend, #152408, Mouse), mCD25 BV605 (clone: PC61, Biolegend, #102035, Mouse), mCD69 PerCpCy5.5 (clone: H1.2F3, Biolegend, #104522, Mouse), mCD3 AF488 (clone: 17A2, Biolegend, #100210, Mouse), mCD4 BUV395 (clone: GK1.5, BD, #563790, Mouse), mCD4 FITC (clone: GK1.5, BD, #553729, Mouse), mCD8 PE-Cy7 (Clone: 53-6.7, Biolegend, #100722, Mouse), hCD87 APC (clone VIM5, eBioscience, #17-3879-42. Human), muPAR PE (R&D Systems, FAB531P. Mouse), muPAR AF700 (R&D Systems, FAB531N. Mouse), huPAR (R&D. AF807. Human), muPAR (R&D. AF534. Mouse), mNRAS (Santa Cruz. SC-31. Mouse, rat, human), mSMA (abcam. Ab5694. Mouse, rat,chicken, piig, cow,dog, human, guinea pig), mKATE (Evrogen. ab233. All species), hCD3 (abcam. ab5690. Human), myc-tag (Cell Signaling. 2276S. All species), mKi-67 (abcam, ab16667.Mouse, rat, human, common marmoset), mIL-6 (abcam. ab6672. Human, monkey), p16-INK4A (Proteintech. 10883-1-AP. Human, monkey), mP-ERKT202/Y204 (Cell Signaling.4370. Human, mouse, rat, hamster, monkey, miink, D.melanogaster, zebrafish, pig, S.cerevisiae), desmin (ThermoFisher Scientific . RB-9014. Mouse, Human). All used antibodies are commercially available.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

ATCC (NALM6, human primary melanocytes). KP cells were a gift from Tyler Jacks. Eu-ALL01 were a gift from Renier J.Brentjens.

Authentication

COA with short tandem repeat was provided with cell line by ATCC. No other authentication was performed. Morphology and properties of all the cell lines pertinent to the experiments (e.g antigen expression or GFP-Luciferase expression) were routinely confirmed by flow cytometry.

Mycoplasma contamination

All cell lines were tested for mycoplasma and were found to be negative.

Commonly misidentified lines
(See ICLAC register)

No commonly misidentified cell lines were used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

NSG (NOD.Cg-Prkdc<scid>Il2rg<tm1Wjl>SzJ) mice male, 8-12 weeks old and obtained from the Jackson Laboratory. C57BL/6N mice were males and females, 8-12 weeks old and obtained from the Jackson Laboratory. B6.SJL-Ptrca/BoyAiTac were females, 6-8 weeks old and obtained from Taconic.

Wild animals

This study did not involve wild animals.

Field-collected samples

This study did not involve samples collected from the field.

Ethics oversight

Memorial Sloan Kettering Cancer Center (MSKCC) Internal Animal Care and Use Committee.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics

Buffy coats from anonymous healthy donors were purchased from the New York Blood Center. The researchers were blind to any covariate characteristics.
Samples from liver fibrosis were obtained from patients with a diagnosis of hepatitis C or B, alcoholic hepatitis or non-alcoholic steatohepatitis from the Biorepository and Pathology CoRE, (Icahn School of Medicine at Mount Sinai).
Samples from normal pancreas and pancreatic tissue with PanIN were obtained from cases with a confirmed diagnosis of pancreatic ductal adenocarcinoma from the Department of Pathology at Memorial Sloan Kettering Cancer Center.
Samples from human carotid sections were obtained from patients undergoing endarterectomy through the Department of Pathology at Weill Medical College of Cornell University.

Recruitment

Buffy coats were purchased from the New York Blood Center. Samples from liver fibrosis were obtained from patients with a confirmed diagnosis of hepatitis C or B, alcoholic hepatitis or non-alcoholic steatohepatitis. Samples were selected by pathologists at the Biorepository and Pathology CoRE, Icahn School of Medicine at Mount Sinai from their archive. Samples from normal pancreas and pancreatic tissue with PanINs were selected from cases with a confirmed diagnosis of pancreatic ductal adenocarcinoma. Samples were selected by pathologists at the Department of Pathology at Memorial Sloan Kettering Cancer Center from their archive. Samples from human carotid sections were obtained from patients undergoing endarterectomy as described in Peerschke, E.I.B. et al. Molecular Immunology. 41; 759-766 (2004): these were all patients with atherosclerotic lesions of type V according to the classification of the American Heart Association. No systematic bias likely to impact results were known at the time of data analysis for any of the samples.

Ethics oversight

All human studies were approved by Mount Sinai, or Weill Medical College of Cornell, or Memorial Sloan-Kettering Institutional Review Board.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Buffy coats from anonymous healthy donors and peripheral blood from healthy volunteers were isolated and purified as described in Methods. For analysis of T cells in the livers, livers were dissociated using MACS Miltenyi Biotec liver dissociation kit (130-1-5-807), filtered through a 100µm strainer, washed with PBS, and red blood cell lysis was achieved with an ACK (Ammonium-Chloride-Potassium) lysing buffer (Lonza). Cells were washed with PBS, resuspended in PEB buffer and used for subsequent analysis. Lungs were minced and digested with 1mg/ml collagenase type IV and DNase type IV in RPMI at 37C and 200rpm for 45 minutes, filtered through 100µm strainer, washed with PBS, and red blood cell lysis was achieved with an ACK lysing buffer (Lonza). Cells were washed with PBS, resuspended in FACS buffer and used for subsequent analysis. For bone marrow samples, tibia and femurs were mechanically disrupted with a mortar in PBS/2mM EDTA, filtered through 40µm strainer, washed with PBS/2mM EDTA and red blood cell lysis was achieved with an ACK lysing buffer (Lonza). Cells were washed with PBS/2mM EDTA, resuspended in FACS buffer and used for subsequent analysis. Spleens were mechanically disrupted with the back of a 5-ml syringe, filtered through 40µm strainer, washed with PBS/2mM EDTA and red blood cell lysis was achieved with an ACK lysing buffer (Lonza). Cells were washed with PBS/2mM EDTA ,resuspended in FACS buffer and used for subsequent analysis.
Cells were subsequently washed, resuspended in FACS Buffer with FcR blocking reagent; antibodies were added and washed off after the incubation time. If intracellular staining was performed, cells were fixed and permeabilized using the Cytofix/Cytoperm kit (BD Biosciences) or Intracellular Fixation & Permeabilizatiion Buffer Set Kit (eBioscience, #88-8824-00) according to the manufacturer's instructions. If cells were counted, counting beads (Invitrogen) were added in the final cell suspension to quantify cells. For analysis of live cells, 7-AAD (BD, #559925, Lot: 9031655, 1:40), Fixable Viability Dye eFluor 506 (65-0866-18, eBioscience, Lot: 2095423, 1:200) and LIVE/DEAD Fixable Violet (L34963, Invitrogen, Lot: 1985351, 1:100) were used.

Instrument

LSRII, BD.
Fortessa 3,BD.
Cytek Aurora (CYTEK)

Software

Collection: FACS DIVA.

Software	Analysis: Flowjo 10.1
Cell population abundance	The purity was verified by flow cytometry.
Gating strategy	The starting cell population was gated on a SSC-A/FSC-A plot. Cell singlets were identified by FSC/SSC gating. Positive/Negative populations were determined by FMO controls.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

DNA of neutrophil extracellular traps promotes cancer metastasis via CCDC25

https://doi.org/10.1038/s41586-020-2394-6

Received: 25 July 2019

Accepted: 1 May 2020

Linbin Yang[1,2], **Qiang Liu**[1,2], **Xiaoqian Zhang**[1,2], **Xinwei Liu**[1,2], **Boxuan Zhou**[1,2], **Jianing Chen**[1,2], **Di Huang**[1,2], **Jiaqian Li**[1,2], **Heliang Li**[1,2], **Fei Chen**[1,2], **Jiang Liu**[1,2], **Yue Xing**[1,2], **Xueman Chen**[1,2], **Shicheng Su**[1,2]✉ & **Erwei Song**[1,2,3,4]✉

Neutrophil extracellular traps (NETs), which consist of chromatin DNA filaments coated with granule proteins, are released by neutrophils to trap microorganisms[1–3]. Recent studies have suggested that the DNA component of NETs (NET-DNA) is associated with cancer metastasis in mouse models[4–6]. However, the functional role and clinical importance of NET-DNA in metastasis in patients with cancer remain unclear. Here we show that NETs are abundant in the liver metastases of patients with breast and colon cancers, and that serum NETs can predict the occurrence of liver metastases in patients with early-stage breast cancer. NET-DNA acts as a chemotactic factor to attract cancer cells, rather than merely acting as a 'trap' for them; in several mouse models, NETs in the liver or lungs were found to attract cancer cells to form distant metastases. We identify the transmembrane protein CCDC25 as a NET-DNA receptor on cancer cells that senses extracellular DNA and subsequently activates the ILK–β-parvin pathway to enhance cell motility. NET-mediated metastasis is abrogated in CCDC25-knockout cells. Clinically, we show that the expression of CCDC25 on primary cancer cells is closely associated with a poor prognosis for patients. Overall, we describe a transmembrane DNA receptor that mediates NET-dependent metastasis, and suggest that targeting CCDC25 could be an appealing therapeutic strategy for the prevention of cancer metastasis.

Recent studies have suggested that NETs, which are released by neutrophils to trap microorganisms during infection, are associated with cancer metastasis in mouse models[5,7]. However, whether NETs are present in human cancers—and if so, their functional role and clinical importance in metastasis—remains largely unknown. Given that NETs trap microorganisms, the proposed mechanisms for their pro-metastatic effects have involved the trapping of disseminated cancer cells[5], but details of the interaction between NETs and cancer cells are not clear.

To investigate the clinical importance of NETs, we conducted immunofluorescence staining for myeloperoxidase (MPO) and citrullinated histone H3 (H3Cit)[8]—specific markers for neutrophils and NETosis, respectively—in the primary tumours and the metastatic lesions of 544 patients with breast cancer. Although NETs were scarce in the primary tumours, they were readily detected in several metastatic lesions, including those in the liver, lungs, bones and brain. Among them, liver metastases exhibited the most abundant NET infiltration (Fig. 1a, b, Extended Data Fig. 1a, b). Notably, serum NET levels—which were evaluated by detecting the MPO–DNA complex as previously reported[9] and showed good agreement with plasma MPO–DNA levels (Extended Data Fig. 1c)—were significantly higher in patients with liver metastases compared with those without metastases or with metastases in other organs (Fig. 1c, Extended Data Fig. 1d). Furthermore, in patients with breast cancer, higher levels of serum MPO–DNA was an independent variable associated with subsequent metastasis to the liver, but not with metastasis to other organs (area under the receiver operating characteristic curve, 0.863; Extended Data Fig. 1e, f). Together, these data implied that excessive amounts of NETs could form in the livers of patients with breast cancer before metastases could be detected, and could facilitate the subsequent development of liver metastases.

To determine whether NETosis in the distant organs precedes cancer metastasis, we inoculated mouse breast cancer 4T1 cells or human breast cancer MDA-MB-231 cells into the mammary fat pads of immunocompetent BALB/c mice or immunocompromised NOD/SCID mice, respectively. In agreement with the clinical findings, positive H3cit signals—indicative of NETs—were evident in the mouse livers but were detected at only very low levels in the lungs (Extended Data Fig. 2a, b). In addition, NETosis in the mouse livers occurred as early as day 16 after the inoculation of human MDA-MB-231 cells—this was much earlier than the detection of hepatic metastases, as assessed by the expression of human *HPRT1* mRNA[10], which occurred at day 34 (Extended Data Fig. 2c). This suggests that hepatic NETs were generated from the infiltrating neutrophils in the livers of tumour-bearing mice before metastases could be detected. Furthermore, we observed that NETs emerged in the livers even before their increase in the plasma on day 25, and earlier than NETosis in the primary tumour tissues on day

---

[1]Guangdong Provincial Key Laboratory of Malignant Tumor Epigenetics and Gene Regulation, Medical Research Center, Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, Guangzhou, China. [2]Breast Tumor Center, Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, Guangzhou, China. [3]Zhongshan School of Medicine, Sun Yat-Sen University, Guangzhou, China. [4]Guangzhou Regenerative Medicine and Health Guangdong Laboratory, Guangzhou, China. ✉e-mail: sushch@mail.sysu.edu.cn; songew@mail.sysu.edu.cn

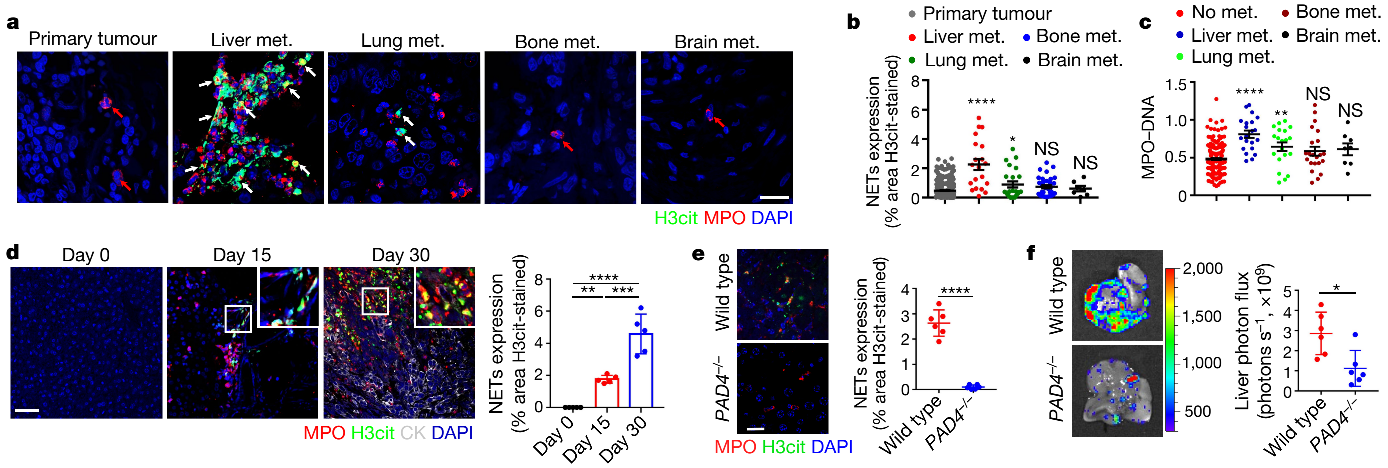


Fig. 1 | NETs in pre-metastatic livers promote cancer metastasis. **a**, Representative immunofluorescence images staining for H3cit and MPO in human primary breast cancer and distant-organ metastases. White arrows indicate NETs, co-stained with H3cit, MPO and DAPI, and red arrows indicate intact neutrophils. Scale bars, 20 μm; met., metastases. **b**, NETs infiltrated in primary breast cancer tissues ($n=461$) and in liver ($n=20$), lung ($n=23$), bone ($n=33$) and brain ($n=7$) metastases. Data are mean ± s.e.m., significance was determined using a two-sided one-way ANOVA with Tukey test, $^{****}P<0.0001$, $^{*}P=0.0204$, NS (not significant) = 0.1508 (bone met.) and 0.9816 (brain met.) compared with primary tumours. **c**, Serum levels of MPO–DNA in patients with breast cancer, without metastases ($n=220$) or with liver ($n=21$), lung ($n=20$), bone ($n=22$) or brain ($n=8$) metastases. Data are mean ± s.e.m., significance was determined using a two-sided one-way ANOVA with Tukey test, $^{****}P<0.0001$, $^{**}P=0.0055$, NS = 0.1430 (bone met.) and 0.3923 (brain met.) compared with patients without metastases. **d**, C57BL/6 mice were intrasplenically injected with E0771 tumour cells, and on the indicated days after injection immunofluorescence images were obtained from the liver tissues, staining for H3cit, MPO and cytokeratin (CK). Scale bars, 50 μm. The white boxes indicate the areas that are magnified in the insets in the top-right corners. $n=5$ biologically independent mice. Data are mean ± s.d., significance was determined using a two-sided one-way ANOVA with Tukey test, $^{**}P=0.0063$, $^{***}P=0.0001$, $^{****}P<0.0001$. **e**, **f**, Representative immunofluorescence images (left) and quantification (right) of NET formation in the liver (**e**) and liver metastases (**f**) after luciferase-E0771 tumour cells (1×10^6) were injected into the spleens of wild-type and *PAD4*$^{-/-}$ female C57BL/6 mice ($n=6$ mice per group). Scale bars, 20 μm. The coloured scale bar represents the intensity of luminescence in arbitrary units. Data are mean ± s.d., significance was determined using a two-tailed Student's *t*-test, $^{****}P<0.0001$, $^{*}P=0.0116$.

37 (Extended Data Fig. 2c), suggesting that the circulating NETs may originate from the livers. To further confirm the association of NETosis with liver metastases, we injected E0771 breast cancer cells into the spleens of C57BL/6 mice and measured NET formation before and after liver metastases could be detected. Abundant NETs were observed in the livers at day 15 after intrasplenic tumour inoculation, earlier than the time at which metastases could be detected (Fig. 1d). Moreover, we observed that NETs in the metastatic livers mainly interacted with the tumour cells (Extended Data Fig. 2d). Collectively, the data suggested that the hepatic metastases of breast cancer were associated with excessive NETosis in the liver.

To investigate the role of NETs in the formation of liver metastases in vivo, we injected E0771 tumour cells into the spleens of C57BL/6 mice deficient in peptidylarginine deiminase 4 (PAD4)—an enzyme required for NET formation[11]—and observed that NET formation and liver metastases were significantly reduced compared with wild-type controls (Fig. 1e, f). To confirm the role of PAD4 in NETosis in our model, we treated the neutrophils isolated from the PAD4-deficient mice with lipopolysaccharide (LPS) and found that in vitro NETosis was significantly inhibited (Extended Data Fig. 2e). To corroborate the knockout mouse model, we treated tumour-bearing mice with DNase I—a nuclease that can degrade NET-DNA[7,12]—and observed that both NETosis and the formation of liver metastases were significantly inhibited (Extended Data Fig. 2f, g). Consistent with the findings in patients with breast cancer, abundant NETs were observed in the liver metastases of patients with colon cancer (Extended Data Fig. 3a, b); furthermore, in the livers of mice that had been intrasplenically injected with HCT116 colon cancer cells, abundant NETs were observed both before and after metastases were detected (Extended Data Fig. 3c, d). Collectively, our data suggest that NETs have an essential role in the cascade by which liver metastases are formed from breast and colon cancers.

To study its pro-metastatic functions, NET-DNA was extracted from neutrophils treated with the NET-inducer phorbol-12-myristate-13-acetate (PMA)[13,14] (Extended Data Fig. 3e). The levels of 8-hydroxy-2′-deoxyguanosine (8-OHdG, a hallmark of NET-DNA[15–17]) were much higher in NET-DNA isolated from the PMA-treated neutrophils than in DNA isolated from untreated neutrophils (Extended Data Fig. 3f, g); in addition, NET-DNA from PMA-treated neutrophils was resistant to TREX-1 (three prime repair exonuclease 1)-mediated degradation (Extended Data Fig. 3h–k). Although previous studies have proposed that NET-DNA serves merely as a trap for 'passer-by' cancer cells[5], our observation that NET-DNA is present both in the blood and in the liver metastases of patients and mouse models implies that NET-DNA—like other chemotactic factors—may attract and interact with the metastatic cancer cells. We therefore evaluated the chemotactic function of NET-DNA using transwell and μ-slide chemotaxis assays. NET-DNA was found to substantially promote the migration and adhesion of MDA-MB-231 cells, an effect that was abrogated upon treatment with DNase I (Extended Data Fig. 3l). In the concentration range 0–5 μg ml^{-1}, NET-DNA enhanced the migration of MDA-MB-231 cells in a dose-dependent manner (Extended Data Fig. 3m); furthermore, MDA-MB-231 cells efficiently migrated towards higher NET-DNA gradient in the chemotactic chamber assay (Extended Data Fig. 3n, Supplementary Video 1). Together, these data indicate that NET-DNA functions as a chemotactic factor to attract cancer cells.

Next, we investigated whether NET-DNA exerts its chemotactic functions by interacting with a DNA receptor on the plasma membrane of cancer cells, as do chemokines. To identify a potential cell-surface DNA receptor, we incubated biotinylated NET-DNA with proteins isolated from the plasma membrane of MDA-MB-231 cells (Extended Data Fig. 4a). Streptavidin beads coupled with biotinylated NET-DNA pulled down a protein of 22 kDa, which was identified as coiled-coil domain containing protein 25 (CCDC25) by liquid chromatography coupled with mass spectrometry (the protein was identified from two tryptic fragments, LHKGENIEDIPK and TADMDVGQIGFHR; Fig. 2a, Extended Data Fig. 4b) and confirmed by immunoblotting using an

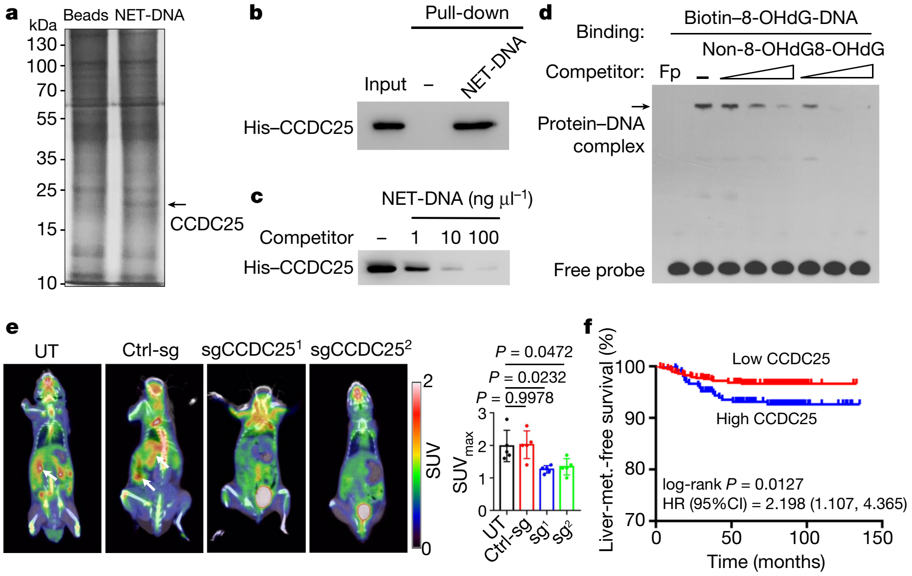


Fig. 2 | NET-DNA binds to CCDC25 on tumour cells and facilitates their distant metastases. a, Purification of NET-DNA-binding proteins. Membrane proteins extracted from MDA-MB-231 tumour cells were incubated either with uncoupled beads or with beads coupled to biotin–NET-DNA. Analysis for bound proteins revealed a 22-kDa protein that bound specifically to NET-DNA (CCDC25, indicated by the arrow). **b**, Purified His-tagged CCDC25 was incubated in the presence or the absence of biotinylated NET-DNA. The bound proteins were immunoprecipitated with streptavidin microbeads and blotted by an anti-His antibody. **c**, Immunoblot analysis for His–CCDC25 immunoprecipitated with biotinylated NET-DNA in the absence or in the presence of increasing concentrations of unbiotinylated NET-DNA. **d**, Competition assays of recombinant CCDC25 bound to 8-OHdG-enriched DNA, competed with 10×, 50× and 100× non-8-OHdG-enriched or 8-OHdG-enriched unlabelled DNA. Fp, free probe. **e**, Mice were intrasplenically injected with MDA-MB-231 cells, which were untreated (UT) or were transduced with a control sgRNA (Ctrl-sgRNA) or with one of two *CCDC25*-targeting sgRNAs (denoted sgCCDC25^1 (or sg^1) and sgCCDC25^2 (or sg^2)). Liver metastases were imaged by ^{18}F-fludeoxyglucose (FDG) positron emission tomography (PET) and computed tomography (CT) and quantified ($n = 5$ mice per group). SUV, standardized uptake value. Data are mean ± s.d., significance was determined using a two-sided one-way ANOVA with Tukey test. **f**, Kaplan–Meier survival curves for patients with breast cancer displaying high ($n = 268$) and low ($n = 573$) CCDC25 expression in the primary tumours. Significance was determined using a two-sided log-rank test. HR, hazard ratio. The data in **a**–**d** are representative of three biologically independent experiments.

anti-CCDC25 antibody (Extended Data Fig. 4c). CCDC25 was then purified from His-tagged CCDC25-overexpressing HEK293T cells, and its binding to biotin-labelled NET-DNA was assessed using an in vitro DNA-binding assay. This confirmed the specific binding of CCDC25 to biotin-labelled NET-DNA (Fig. 2b), which was efficiently competed by the unlabelled NET-DNA in a dose-dependent manner (Fig. 2c). Mixing the cytoplasmic membrane proteins of MDA-MB-231 cells with NET-DNA and running an electrophoretic mobility shift assay (EMSA) revealed the formation of a protein–DNA complex, which could be super-shifted by the anti-CCDC25 antibody (Extended Data Fig. 4d). In addition, CCDC25 knockout in the MDA-MB-231 cells abrogated the formation of the protein–DNA complex (Extended Data Fig. 4e). The apparent dissociation constant (K_d) of recombinant CCDC25 for isolated NET-DNA, as determined by EMSA assays, was 67.24 ± 5.09 nM (mean ± s.d.; Extended Data Fig. 4f, g), suggesting highly efficient binding of CCDC25 to NET-DNA. Furthermore, pretreatment of the isolated NET-DNA with DNase I—but not with Proteinase K—abrogated the interaction between CCDC25 and NET-DNA (Extended Data Fig. 4h), indicating that CCDC25 binds to DNA but not to the DNA-associated protein. This was further validated by in vitro binding assays showing that recombinant CCDC25 efficiently bound to synthetic DNA, forming stronger interactions with 8-OHdG-enriched DNA than with non-8-OHdG-enriched DNA (Extended Data Fig. 4i, j). To confirm this preference, we used bio-layer interferometry[18] to measure the binding capacities of CCDC25 with 8-OHdG-enriched DNA and with non-8-OHdG-enriched DNA. CCDC25 exhibited a 4.4-fold higher affinity for 8-OHdG-enriched DNA ($K_d = 6.0 \pm 1.5$ nM) than for non-8-OHdG-enriched DNA ($K_d = 26.6 \pm 6.0$ nM) (Extended Data Fig. 4k). We further confirmed the specificity of CCDC25 for 8-OHdG-enriched DNA by performing competition assays, and observed that CCDC25 bound to 8-OHdG-enriched biotin-labelled DNA was competed more efficiently by the 8-OHdG-enriched non-labelled DNA than by the non-8-OHdG-enriched non-labelled DNA (Fig. 2d). Taken together, these data reveal that CCDC25 serves as a sensor for 8-OHdG-enriched NET-DNA on the plasma membrane.

In vitro, CCDC25 knockout in MDA-MB-231 cells efficiently abolished their adhesion to culture plates coated with NET-DNA, but not to those coated with fibronectin (Extended Data Fig. 5a, b). Moreover, CCDC25 knockout abrogated cytoskeleton remodelling of cancer cells induced by NET-DNA (Extended Data Fig. 5c) and reduced their chemotaxis towards NET-DNA (Extended Data Fig. 5d, Supplementary Video 2). We also found that CCDC25 knockout moderately inhibited the proliferation of cancer cells induced by NET-DNA, but not by a panel of proliferation-inducing cytokines (Extended Data Fig. 5e). These results indicate that NET-DNA induces migration, adhesion and proliferation of tumour cells via interaction with CCDC25.

To evaluate whether CCDC25 also mediates a tumour-cell response to the DNA of apoptotic cells, we treated wild-type and CCDC25-knockout MDA-MB-231 cells with apoptotic DNA isolated from tumour cells exposed to docetaxel[19]. DNA derived from the apoptotic cells slightly enhanced the migration and adhesion of the wild-type MDA-MB-231 cells, but to a much weaker extent compared with NET-DNA. Moreover, CCDC25 depletion abrogated the response of tumour cells towards apoptotic DNA (Extended Data Fig. 5f). Consistent with our findings that CCDC25 preferentially binds to 8-OHdG DNA, the level of 8-OHdG modification was found to be much higher in NET-DNA than in the apoptotic DNA (Extended Data Fig. 5g). Collectively, these data suggest that CCDC25 mediates the chemotactic response of cancer cells to extracellular DNAs, especially to those with 8-OHdG modification.

We next investigated the role of CCDC25 in NET-mediated metastasis in vivo. Knocking out CCDC25 in the MDA-MB-231 cells that were

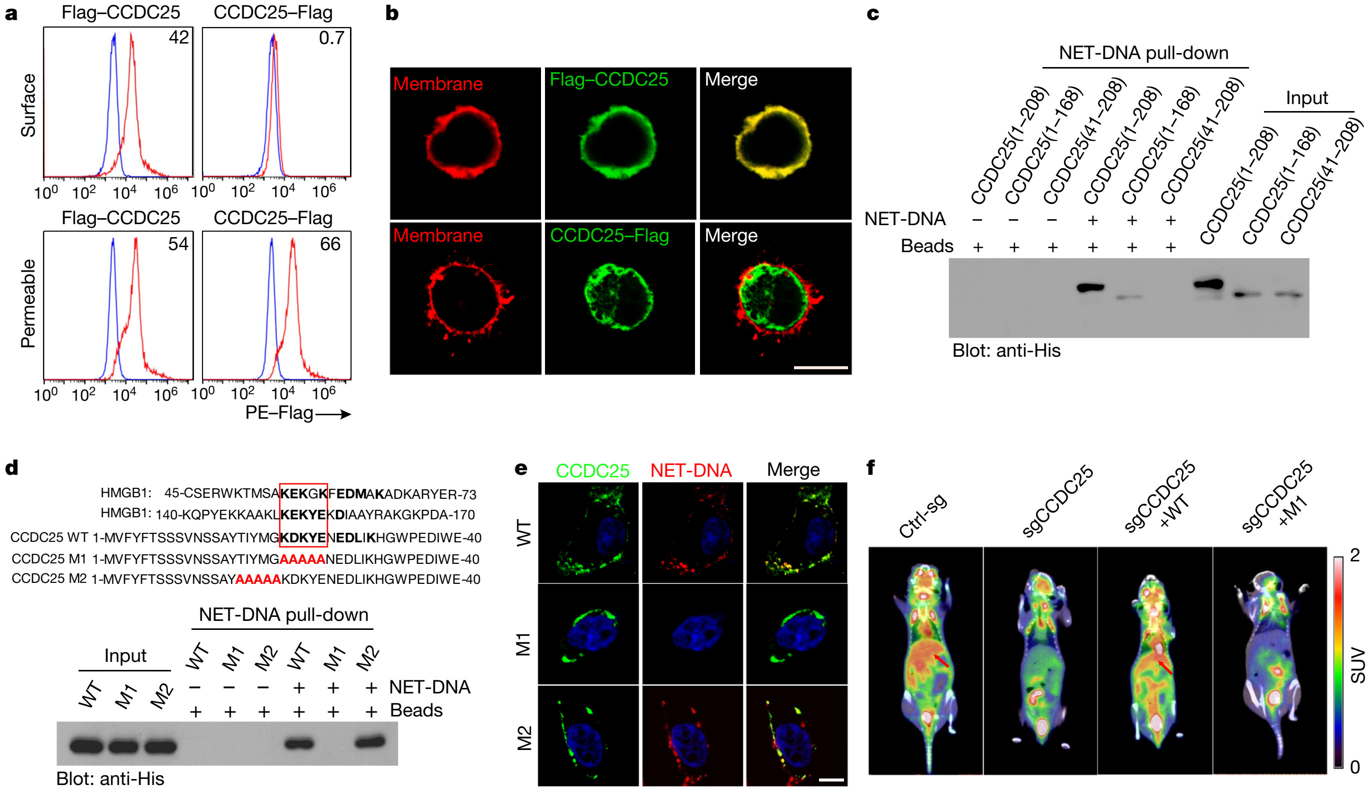


Fig. 3 | The transmembrane protein CCDC25 interacts with NET-DNA at its N terminus. a, Flow cytometry analysis of Flag signals in HeLa cells with ectopic expression of CCDC25 tagged with Flag at the N terminus (left, Flag–CCDC25) or the C terminus (right, CCDC25–Flag). Surface, surface staining without permeabilization; Permeable, cells were permeabilized before staining. The numbers in the histograms indicate the mean fluorescence intensity (×10^3) of anti-Flag staining. **b**, Confocal microscopy images of the ectopically expressed Flag–CCDC25 (top) or CCDC25–Flag (bottom). The cytoplasmic membrane was stained with Dil (red) and CCDC25 with anti-Flag (green). Scale bars, 10 μm. **c**, The interaction between His-tagged full-length or truncated CCDC25 and NET-DNA was analysed by precipitation of NET-DNA followed by immunoblotting using anti-His antibody. **d**, Sequence alignment of the CCDC25 N terminus with two DNA-binding domains of a classical DNA sensor (HMGB1). The interaction between His-tagged full-length CCDC25 (wild-type, WT), the AA_{21-25} mutant (M1) or the AA_{16-20} mutant (M2) with NET-DNA was analysed by precipitation of NET-DNA followed by immunoblotting using anti-His antibody. **e**, Confocal microscopy images showing the co-localization of NET-DNA with CCDC25. HeLa cells with ectopic expression of eGFP-tagged full-length or mutated CCDC25 were incubated with NET-DNA stained with SYTOX Orange nucleic acid (red). NET-DNA (red) co-localized with full-length CCDC25 (green) and the AA_{16-20} mutant (M2), but not with the AA_{21-25} mutant (M1). Scale bars, 10 μm. **f**, CCDC25-knockout MDA-MB-231 cells with or without ectopic expression of full-length CCDC25 (WT) or the AA_{21-25} mutant (M1) were injected intrasplenically into NOD/SCID mice. Liver metastases were visualized by ^{18}F-FDG PET/CT imaging ($n = 5$ mice per group). Data in **a**–**e** are representative of three biologically independent experiments.

inoculated into NOD/SCID mice significantly reduced NET-mediated lung metastases induced by nasal instillation of LPS[7] (Extended Data Fig. 5h), and inhibited the formation of liver metastases upon intrasplenic injection of the cells into the mice without pretreatment (Fig. 2e). Moreover, enforced expression of CCDC25 in MCF-7 cancer cells promoted their adhesion and migration towards NET-DNA in vitro, and significantly enhanced formation of their lung and liver metastases in vivo (Extended Data Fig. 5i–l, Supplementary Video 3). Similar results were observed for HCT116 colon cancer cells (Extended Data Fig. 5m–q, Supplementary Video 4).

To investigate the role of CCDC25 in more clinically relevant models, we isolated primary cancer cells from patients with breast cancer and depleted the cells of the encoding gene, *CCDC25*, using CRISPR–Cas9 techniques. We found that CCDC25 knockout significantly abolished adhesion, migration and cytoskeleton remodelling of primary cancer cells towards NETs in vitro (Extended Data Fig. 6a) and inhibited liver metastases of the primary tumour cells that were intrasplenically injected into NOD/SCID mice (Extended Data Fig. 6b). Furthermore, we generated CCDC25-knockout mice (*CCDC25*$^{-/-}$) and crossed them with MMTV-PyMT mice to generate PyMT;CCDC25$^{-/-}$ hybrids. We found that CCDC25 knockout markedly reduced lung metastases induced by nasal instillation of LPS (Extended Data Fig. 6c) but did not affect primary tumour growth (Extended Data Fig. 6d). In addition, we injected the breast tumour cells isolated from PyMT and PyMT;CCDC25$^{-/-}$ mice into the spleens of syngeneic C57BL/6 mice and observed that CCDC25 depletion significantly inhibited liver metastases (Extended Data Fig. 6e). We then generated a polyclonal antibody against CCDC25 (Extended Data Fig. 6f), which efficiently bound to recombinant CCDC25 (Extended Data Fig. 6g). In vitro, the CCDC25 antibody effectively inhibited the migration, adhesion and cytoskeleton remodelling of cancer cells induced by NETs (Extended Data Fig. 6h). More notably, we observed that the CCDC25 antibody markedly inhibited the formation of liver metastases when MDA-MB-231 cells were injected into the spleens of NOD/SCID mice (Extended Data Fig. 6i).

Clinically, CCDC25 expression has been noted in several types of cancer, with high expression levels reported in the breast and colon cancers indexed in the Human Protein Atlas database (Extended Data Fig. 7a). In our cohorts of patients with breast cancer and patients with colon cancer, CCDC25 expression was detected in both the cytoplasm and the cytoplasmic membrane of the tumour cells (Extended Data Fig. 7b–d), and we observed a clear membrane staining of CCDC25 at the invasive front of the tumours (Extended Data Fig. 7e). CCDC25 also closely interacted with NET-deposited H3cit in the metastatic liver tumours of patients with breast cancer (Extended Data Fig. 7f). Notably, higher levels of CCDC25 expression in the primary tumours was associated with reduced long-term survival (Fig. 2f, Extended Data

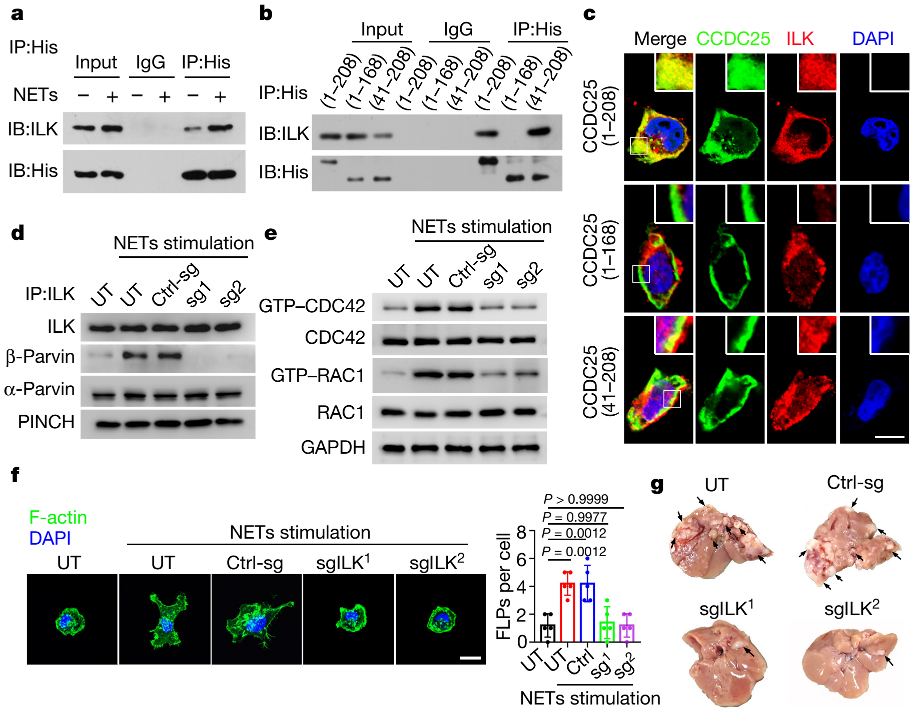


Fig. 4 | CCDC25 interacts with ILK at its C terminus and signals through the ILK–β-parvin cascade. a, Immunoblotting (IB) of ILK (top) or His–CCDC25 (bottom) in the lysates (input) or immunoprecipitates (IP; IgG or anti-His) of HeLa cells transfected with His-tagged CCDC25 with (+) or without (−) NETs treatment. **b**, Immunoblotting of ILK (top) or His–CCDC25 (bottom) in the lysates (input) or immunoprecipitates (IgG or anti-His) of HeLa cells transfected with His-tagged full-length or truncated CCDC25. **c**, Confocal microscopy images showing the colocalization of CCDC25 with ILK in HeLa cells transfected with His-tagged full-length or truncated CCDC25. Scale bars, 10 μm. **d**, Immunoblotting for the indicated proteins in the anti-ILK immunoprecipitates from the lysates of MDA-MB-231 cells either untreated or transduced with control or *CCDC25* sgRNAs, with or without NETs stimulation. **e**, GTP-bound or total RAC1 and CDC42 levels were examined in the lysates of MDA-MB-231 cells, either untreated or transduced with control or CCDC25 sgRNAs and stimulated with or without NETs. **f**, MDA-MB-231 cells untreated or transduced with a control or with one of two *ILK* sgRNAs were stimulated with or without NETs, and then stained with phalloidin (F-actin, green). FLPs, filopodium-like protrusions. Scale bars, 10 μm. $n = 5$ biologically independent experiments. Data are mean ± s.d., two-sided one-way ANOVA with Tukey test. **g**, Representative images of liver metastases in NOD/SCID mice 30 days after intrasplenic injection of MDA-MB-231 cells, which were untreated or were transduced with a control sgRNA or with one of two *ILK* sgRNAs. $n = 6$ mice per group. Data in **a**–**e** are representative of three biologically independent experiments.

Fig. 7g, h); this was further corroborated for various types of cancer in several online databases (Extended Data Fig. 7i).

We next investigated how CCDC25 interacts with NET-DNA. Transmembrane helix prediction algorithms predicted CCDC25 to be a single transmembrane protein, with a hydrophobic centre between residues 60 and 80 (hydrophobicity score greater than 0) and two hydrophilic ends (Extended Data Fig. 8a). To validate the identity of CCDC25 as a transmembrane protein, we used a membrane fractionation assay as previously reported[20]. CCDC25 was precipitated in the membranous fractions at high salt concentrations, high pH or in the presence of a strong denaturant, but was solubilized into the supernatant by Triton X-100 (Extended Data Fig. 8b). Additionally, flow cytometry analysis demonstrated that CCDC25 with a Flag tag at the N terminus could be stained with or without permeabilizing the cells, whereas CCDC25 with a Flag tag at the C terminus could be stained only in permeabilized cells (Fig. 3a). These findings suggest that CCDC25 on the plasma membrane consists of an extracellular N terminus and an intracellular C terminus; this was further confirmed by confocal microscopy using immunofluorescence co-staining with the DiI and Flag antibodies (Fig. 3b). Furthermore, deletion of 40 amino acids at the N terminus—but not the C-terminal 40 amino acids (Extended Data Fig. 8c)—completely abolished the interaction between CCDC25 and NET-DNA (Fig. 3c). To identify the exact DNA-binding domain at the N terminus of CCDC25, we compared its amino acid sequence with that of several DNA sensors. We found that amino acids 21–25 (AA_{21-25}) of CCDC25 (KDKYE) were aligned with the DNA-binding domain of high mobility group box 1 (HMGB1). Mutation of AA_{21-25}, but not of the adjacent AA_{16-20}, abrogated the binding of CCDC25 to NET-DNA (Fig. 3d, e, Extended Data Fig. 8d). Additionally, ectopic expression of wild-type CCDC25, but not of its AA_{21-25} mutant, rescued the NET-DNA-induced cytoskeleton rearrangement and directional chemotaxis in CCDC25-knockout MDA-MB-231 cells (Extended Data Fig. 8e–g, Supplementary Video 5). In vivo, overexpression of the wild-type CCDC25, but not of its AA_{21-25} mutant, in the MDA-MB-231 cells inoculated into mouse spleens was found to reverse the inhibition of liver metastases mediated by CCDC25 knockout (Fig. 3f, Extended Data Fig. 8h). These data suggest that the AA_{21-25} domain at the extracellular N terminus of CCDC25 is the binding site for NET-DNA.

To explore the downstream signalling of CCDC25, we performed a pull-down assay using the lysate of HeLa cells transfected with His-tagged CCDC25 and analysed the proteins that bound to CCDC25 by mass spectrometry. We identified integrin-linked kinase (ILK) as interacting with CCDC25 in the cells stimulated with NET-DNA (Extended Data Fig. 9a–c); this was validated by an immunoprecipitation assay and by confocal microscopy (Fig. 4a, Extended Data Fig. 9d). Endogenous CCDC25 also bound to ILK, and this binding was further enhanced upon stimulation by NETs (Extended Data Fig. 9e). Furthermore, deletion of the intracellular C terminus of CCDC25—but not the extracellular N terminus—abolished its interaction with ILK (Fig. 4b–c). These findings indicate that the intracellular C terminus of CCDC25 interacts with ILK after stimulation by NET-DNA.

Using an immunoprecipitation assay, we found that ILK recruited β-parvin—but not α-parvin or PINCH (particularly interesting new cysteine-histidine-rich protein)—after stimulation by NET-DNA. This recruitment was attenuated upon CCDC25 knockout (Fig. 4d). Furthermore, NET-DNA stimulation significantly increased the levels of GTP-bound RAC1 and CDC42 (Fig. 4e) but had no appreciable effects on the phosphorylation of AKT or GSK3β (Extended Data Fig. 10a). The knockout of CCDC25 and ILK or the silencing of β-parvin substantially reduced the activation of RAC1 and CDC42 (Fig. 4e, Extended Data Fig. 10b, c), and markedly inhibited NET-DNA-induced actin remodelling and directional chemotaxis in MDA-MB-231 cancer cells (Fig. 4f, Extended Data Fig. 10d–f, Supplementary Videos 6, 7). ILK knockout also abolished the NET-induced proliferation of tumour cells (Extended Data Fig. 10g, h). In vivo, significantly fewer liver metastases were observed after intrasplenic injection of ILK-knockout MDA-MB-231 cells (Fig. 4g, Extended Data Fig. 10i). Together, these data suggest that NET-DNA binds to CCDC25 and triggers the ILK–β-parvin–RAC1–CDC42 cascade in order to promote the metastasis of cancer cells.

This study identifies CCDC25—a transmembrane protein on the cytoplasmic membrane—as a specific sensor for DNA, in particular for the 8-OHdG-enriched DNA present in NETs. After sensing NET-DNA at AA_{21-25} on its extracellular domain, CCDC25 recruits ILK via its intracellular C terminus and initiates the β-parvin–RAC1–CDC42 cascade to induce cytoskeleton rearrangement and directional migration of tumour cells (Extended Data Fig. 10j). Clinically, NETosis is abundant in the liver metastases of patients with breast cancer and patients with colon cancer. Increased levels of NETs in blood could act as a biomarker to specifically predict the long-term risk of liver metastases in patients with early-stage breast cancer. In mouse models in vivo, the targeting of CCDC25 reduced the formation of NET-mediated distant metastases. In summary, we show that an extracellular DNA sensor located on the cytoplasmic membrane could potentially serve as a therapeutic target for metastasis.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2394-6.

1. Brinkmann, V. et al. Neutrophil extracellular traps kill bacteria. *Science* **303**, 1532–1535 (2004).
2. McDonald, B., Urrutia, R., Yipp, B. G., Jenne, C. N. & Kubes, P. Intravascular neutrophil extracellular traps capture bacteria from the bloodstream during sepsis. *Cell Host Microbe* **12**, 324–333 (2012).
3. Fuchs, T. A. et al. Novel cell death program leads to neutrophil extracellular traps. *J. Cell Biol*. **176**, 231–241 (2007).
4. Tohme, S. et al. Neutrophil extracellular traps promote the development and progression of liver metastases after surgical stress. *Cancer Res*. **76**, 1367–1380 (2016).
5. Cools-Lartigue, J. et al. Neutrophil extracellular traps sequester circulating tumor cells and promote metastasis. *J. Clin. Invest*. **123**, 3446–3458 (2013).
6. Cedervall, J., Zhang, Y. & Olsson, A. K. Tumor-induced NETosis as a risk factor for metastasis and organ failure. *Cancer Res*. **76**, 4311–4315 (2016).
7. Albrengues, J. et al. Neutrophil extracellular traps produced during inflammation awaken dormant cancer cells in mice. *Science* **361**, eaao4227 (2018).
8. Park, J. et al. Cancer cells induce metastasis-supporting neutrophil extracellular DNA traps. *Sci. Transl. Med*. **8**, 361ra138 (2016).
9. Kessenbrock, K. et al. Netting neutrophils in autoimmune small-vessel vasculitis. *Nat. Med*. **15**, 623–625 (2009).
10. Yu, F. et al. *let-7* regulates self renewal and tumorigenicity of breast cancer cells. *Cell* **131**, 1109–1123 (2007).
11. Li, P. et al. PAD4 is essential for antibacterial innate immunity mediated by neutrophil extracellular traps. *J. Exp. Med*. **207**, 1853–1862 (2010).
12. van der Windt, D. J. et al. Neutrophil extracellular traps promote inflammation and development of hepatocellular carcinoma in nonalcoholic steatohepatitis. *Hepatology* **68**, 1347–1360 (2018).
13. Noubouossie, D. F. et al. In vitro activation of coagulation by human neutrophil DNA and histone proteins but not neutrophil extracellular traps. *Blood* **129**, 1021–1029 (2017).
14. Warnatsch, A., Ioannou, M., Wang, Q. & Papayannopoulos, V. Neutrophil extracellular traps license macrophages for cytokine production in atherosclerosis. *Science* **349**, 316–320 (2015).
15. Yang, H. et al. New insights into neutrophil extracellular traps: mechanisms of formation and role in inflammation. *Front. Immunol*. **7**, 302 (2016).
16. Lood, C. et al. Neutrophil extracellular traps enriched in oxidized mitochondrial DNA are interferogenic and contribute to lupus-like disease. *Nat. Med*. **22**, 146–153 (2016).
17. Gehrke, N. et al. Oxidative damage of DNA confers resistance to cytosolic nuclease TREX1 degradation and potentiates STING-dependent immune sensing. *Immunity* **39**, 482–495 (2013).
18. Berlow, R. B., Dyson, H. J. & Wright, P. E. Hypersensitive termination of the hypoxic response by a disordered protein switch. *Nature* **543**, 447–451 (2017).
19. Chen, F. et al. Extracellular vesicle-packaged HIF-1α-stabilizing lncRNA from tumour-associated macrophages regulates aerobic glycolysis of breast cancer cells. *Nat. Cell Biol*. **21**, 498–510 (2019).
20. Schuberth, C. & Buchberger, A. Membrane-bound Ubx2 recruits Cdc48 to ubiquitin ligases and their substrates to ensure efficient ER-associated protein degradation. *Nat. Cell Biol*. **7**, 999–1006 (2005).

Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment, except for when noted otherwise.

Cell lines and neutrophil isolation

Human breast cancer cell lines MDA-MB-231 and MCF-7, human colon cancer cell line HCT116, mouse breast cancer cell line 4T1, HEK293T cells and HeLa cells were purchased from ATCC, and the mouse E0771 breast cancer cell line was purchased from CH3 Biosystems. The cell lines were authenticated by short tandem repeat profiling and were tested negative for mycoplasma contamination. Human neutrophils were isolated from the peripheral blood of healthy donors using Ficoll density gradient centrifugation and isolated by positive selection for CD66b$^+$ cells with microbeads (Miltenyi Biotec, 130-104-913). All the cells were cultured in a humidified, 5% CO_2 incubator at 37 °C, and grown in RPMI or DMEM with 10% fetal bovine serum.

Patients and tissue samples

Immunofluorescence staining for NETs was performed in the tissues of primary tumours (461 cases), liver metastases (20 cases), lung metastases (23 cases), bone metastases (33 cases) and brain metastases (7 cases) of patients with breast cancer, collected from Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University (Guangzhou, China) between 2007 and 2016. The serum and plasma samples were collected from patients with breast cancer who were admitted to Sun Yat-Sen Memorial Hospital between 2011 and 2019. Moreover, immunofluorescence staining for NETs was also performed in the tissues of primary tumours (130 cases), liver metastases (16 cases), lung metastases (12 cases), bone metastases (3 cases) and brain metastases (5 cases) of patients with colon cancer who were admitted to Sun Yat-Sen Memorial Hospital. Immunohistochemical staining for CCDC25 was performed in the primary tumour tissues (202 cases) and liver metastases (17 cases) of patients with breast cancer and in the primary tumour tissues (134 cases) and liver metastases (16 cases) of patients with colon cancer. Additionally, immunohistochemical staining for CCDC25 was performed in 841 breast cancer samples and 134 colon cancer samples from Sun Yat-Sen Memorial Hospital. All samples were collected from patients who had provided informed consent, and all the related procedures were performed with the approval from the internal review and ethics board of Sun-Yat-Sen Memorial Hospital.

Immunofluorescence staining

Tissue was fixed in 4% paraformaldehyde (Thermo Scientific) for 24 h at 4 °C, washed with PBS, embedded in paraffin and sectioned at 4-µm thickness. Antigen retrieval was performed using target retrieval solution, pH 9.0 (Dako) in a pressure cooker for 15–20 min. Non-specific binding was then blocked with 5% BSA for 25 min at room temperature. Cells for immunofluorescence were fixed with 4% paraformaldehyde for 25 min at room temperature, washed with PBS and permeabilized with or without 0.2% Triton X-100 in PBS for 20 min. Cells were then blocked in PBS with 2% BSA for 30 min at room temperature. Subsequently, the samples were incubated with rabbit anti-H3Cit (1:100, Abcam, ab5103), goat anti-MPO (10 µg ml^{-1}, R&D, AF3667), mouse anti-cytokeratin (1:50, GeneTex, GTX27753), mouse anti-8-OHdG (1:50, GeneTex, GTX41980), rabbit anti-Flag (1:100, CST, 14793), rabbit anti-ILK (1:100, CST, 3862), mouse anti-His tag (1:100, Thermo Scientific, MA1-21315), rabbit anti-CCDC25 (1:50, Invitrogen, PA5-54735), mouse anti-CCDC25 (1:50, Santa Cruz, sc-515201), rabbit anti-Ki67 (1:50, Abcam, ab16667), mouse anti-CD31 (1:50, Abcam, ab9498), rat anti-CD31 (1:50, Abcam, ab56299), mouse anti-PDGFRβ (1:50, Abcam, ab69506) or mouse anti-α-SMA (1:50, R&D, MAB1420) overnight at 4 °C. The tissues were incubated with Alexa-Fluor-conjugated secondary antibodies (Invitrogen) in 1% BSA for 1 h at room temperature. Filamentous actin (F-actin) was stained with Alexa Fluor 488 Phalloidin (165 nM, A12379, Invitrogen) or Alexa Fluor 555 Phalloidin (165 nM, A34055, Invitrogen) at room temperature for 20 min. The plasma membrane was labelled with CM-Dil (1 µM, C7000, Thermo Scientific) at room temperature for 10 min and then for an additional 10 min at 4 °C. DAPI was then used for counterstaining the nuclei and images were obtained by laser scanning confocal microscopy (LSM800, Zeiss). For the visualization of NETs in vitro, mouse neutrophils were stimulated with 1 µg ml^{-1} LPS for 3 h (ref. [11]). The NETs were determined as the percentage of the positive H3cit signal[8,21] in each field of view in the overall tissues by Imaris 9.0 Microscopy Image Analysis Software. For NET quantification, NETs were counted in at least 10 fields per section and 5 sections per sample were evaluated. The accuracy of automated measurements was confirmed by two independent investigators (D.H. and F.C.), who were unaware of the patients' clinical information.

Detection of serum MPO–DNA

We detected serum MPO–DNA using a previously described capture ELISA method with slight modifications[9,22]. 96-well microtiter plates were coated with 5 µg ml^{-1} anti-MPO monoclonal antibody (ABD Serotec, 0400-0002) as the capturing antibody overnight at 4 °C. After blocking in 1% BSA, patient serum together with peroxidase-labelled anti-DNA monoclonal antibody was added (component No.2 of the Cell Death Detection ELISA kit, Roche, 11774425001), incubated at room temperature for 2 h and then washed with PBS three times. The peroxidase substrate (Roche, 11774425001) was added. After incubation at 37 °C for 40 min, the optical density was measured at 405 nm using a microplate reader (infinite M200 PRO).

Animal experiments

PAD4$^{flox/flox}$ mice (B6(Cg)-*Padi4$^{tm1.2Kmow}$*/J, 026708) and Ddx4-Cre mice (B6.FVB-Tg(Ddx4-cre)1Dcas/KnwJ, 018980) were purchased from the Jackson Laboratory. *PAD4$^{flox/flox}$* mice were crossed with Ddx4-Cre mice to generate *PAD4$^{+/-}$* mice, which were then intercrossed to generate *PAD4$^{-/-}$* mice and bred in the specific-pathogen-free animal facility of the Animal Experiment Center of Sun-Yat-Sen University. The *CCDC25$^{-/-}$* mice in a C57BL/6 genetic background were generated by deleting the genomic DNA fragment covering exon 3 using a CRISPR–Cas9-mediated genome editing system by Shanghai Model Organisms Center. B6.FVB-Tg (MMTV-PyMT) (C57BL/6 background) mice were purchased from the Jackson Laboratory and were crossed with *CCDC25$^{-/-}$* mice to obtain CCDC25-knockout spontaneously tumorigenic mice (PyMT;CCDC25$^{-/-}$). Six- to eight-week-old female C57BL/6 mice, BALB/c mice and NOD/SCID mice were maintained at the Animal Experiment Center of Sun-Yat-Sen University, and all procedures were approved by the Animal Care and Use Committee of Sun-Yat-Sen University. Mice were randomized at the beginning of each experiment and experiments were not blinded. For orthotopic transplantations, 1×10^6 MDA-MB-231 or 5×10^5 4T1 cells were re-suspended in 100 µl of PBS and injected in the fourth mammary fat pads on one flank of the mice. At various time points (noted in the relevant figure legends) after tumour inoculation, the mice were euthanized. The maximally permitted tumour diameter of 20 mm in any dimension was never exceeded. For experimental liver metastases, 1×10^6 MDA-MB-231-luc and primary breast cancer cells transduced with a control single-guide RNA (sgRNA) or two sgRNAs for *CCDC25*, or 1×10^6 HCT116-luc tumour cells transduced with or without lentivirus carrying control or *CCDC25* shRNA, 1×10^6 E0771-luc cells, 2×10^6 MCF-7-luc with or without *CCDC25* overexpression were resuspended in 50 µl of PBS and intrasplenically injected. The mice were euthanized 30 days after injection. In some experiments, 50 µl PBS with or without 0.25 mg ml^{-1} LPS was administered intranasally on days 0, 3 and 6 with a P200 pipette into mice under anaesthesia (2.5% isoflurane)[7]. Breast cancer MDA-MB-231-luc cells (1×10^6) transduced with a control sgRNA or sgRNA for *CCDC25*, colon cancer HCT116-luc cells (1×10^6) transduced with lentivirus carrying control or *CCDC25* shRNA and MCF-7-luc (2×10^6) with or without *CCDC25* overexpression were resuspended in 100 µl of PBS and injected into the tail vein. The mice were euthanized 30 days after injection. For the LPS-induced lung inflammation in MMTV-PyMT and PyMT;CCDC25$^{-/-}$ transgenic tumour models, 50 µl PBS with or without 0.25 mg ml^{-1} LPS was administered intranasally with a P200 pipette into mice under anaesthesia (2.5% isoflurane), every 3 days for a total of 5 times, starting at 18 weeks of age. The tumour weight measurements and analysis of lung metastases were performed at 24 weeks of age. We examined the metastases using a PET/CT imaging system (Siemens) or a IVIS Lumina Imaging System (Xenogen). The liver and lung tissues were collected for further evaluation.

Primary cancer cell culture

Primary breast cancer cells were isolated from invasive ductal carcinoma samples obtained from surgery. In brief, the tissues were digested using collagenase type I, collagenase type III and hyaluronidase (1.5 mg ml^{-1}, Sigma Aldrich) at 37 °C with agitation for 2–3 h in DMEM with 10% FBS. Single-cell suspensions were obtained by filtration through a 40-µm filter and the primary cancer cells were purified by magnetic-activated cell sorting with CD326 (EpCAM) Tumour Cell Enrichment and Detection Kit (Miltenyi Biotec, 130-090-500) according to the manufacturer's instructions[23].

Intrasplenic MMTV-PyMT tumour model

An intrasplenic MMTV-PyMT tumour model was obtained by transplanting tumour-derived cells from transgenic MMTV-PyMT or PyMT;CCDC25$^{-/-}$ mice into the spleen of C57BL/6 mice[24]. In brief, late-stage mammary carcinomas of 16-week-old female MMTV-PyMT or PyMT;CCDC25$^{-/-}$ mice were digested using collagenase type I, collagenase type III and hyaluronidase (1.5 mg ml^{-1}, Sigma Aldrich) at 37 °C with agitation for 1 h in DMEM medium. The cell suspensions were then filtered through a 40-µm filter and washed with PBS 3 times. The resultant cells (1×10^6) were injected into the spleen of a recipient mouse. Thirty days after tumour injection, mice were euthanized and the metastatic liver burden was measured.

Isolation of cell membrane protein

We isolated cell membrane protein with Pierce Cell Surface Protein Isolation Kit (Thermo Scientific, 89881) as previously reported[25]. In brief, the cells were labelled with Thermo Scientific EZ-Link Sulfo-NHS-SS-Biotin, which is a non-cell-permeable and thiol-cleavable amine-reactive biotinylation reagent. Thereafter, the cells were lysed with Pierce IP Lysis Buffer (25 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% NP-40 and 5% glycerol) and the biotinylated cell membrane proteins were purified by NeutrAvidin-agarose resin. The cell membrane proteins were released by incubation with Pierce IP Lysis Buffer containing 5 mM dithiothreitol.

Purification of NETs

We isolated NETs from primary human neutrophils of peripheral blood using a previously described method with slight modifications[13,14]. Neutrophils were treated with 500 nM PMA for 4 h. After removal of the supernatant, NETs adhered at the bottom were washed down by pipetting 2 ml of cold PBS and were centrifuged at 1,000*g* at 4 °C for 10 min. The cell-free supernatant containing NETs (DNA–protein complex) was collected. The DNA concentration of NETs was measured by spectrophotometry and the NETs were used for further experiments.

Scanning electron microscopy

Neutrophils were added on the coverslips and either treated with 500 nM PMA for 4 h or left untreated. To detect the cell-free NETs, the aforementioned isolated NETs were added and coated on coverslips at 37 °C overnight. Thereafter, the samples were processed for scanning electron microscopy as previously described[1]. In brief, the samples were fixed with 2.5% glutaraldehyde overnight. After fixation, the samples were washed with PBS, incubated with 1% osmium tetroxide and dehydrated with a graded ethanol series, followed by critical-point drying and coating with 2 nm platinum. Next, the samples were coated with 5 nm carbon and analysed using an FEI Quanta 200 scanning electron microscope.

Purification of NET-DNA

The aforementioned isolated NETs were fragmented with a VCX130 sonicator to the length of 200–500 bp. The NET-DNA was purified using a MicroElute DNA Clean Up Kit (OMEGA, D6296) and biotinylated using Biotin 3′End DNA Labelling Kit (Thermo Scientific, 89818) according to the manufacturer's instructions.

Apoptotic DNA isolation

MDA-MB-231 cells were treated with docetaxel (2 µg ml^{-1}) for 12 h (ref. [19]), dissociated with 0.25% trypsin-EDTA and collected by centrifugation. The apoptotic DNA was extracted with an apoptotic DNA purification kit according to the manufacturer's instructions (Beyotime, C0008).

Recombinant protein production and purification

Full-length CCDC25 was cloned into pET28a expression vector (Novagen) with a His fusion at the C terminus. The plasmid was transformed into *Escherichia coli* BL21(DE3), inoculated in LB medium (containing kanamycin) and grown at 37 °C until an optical density at 600 nm (OD_{600}) of 0.7 was reached, then isopropyl-β-D-thiogalactopyranoside (to a final concentration of 0.4 mM) was added. After growing for 3 h at 37 °C, cells were pelleted and lysed in buffer (20 mM sodium phosphate, pH 7.5, 10 mM imidazole, 0.5 M NaCl and EDTA-free protease inhibitors). The mixture was sonicated and the insoluble protein was removed by centrifugation. The His-tagged protein was isolated from the supernatants using a HisPur Ni-NTA Purification Kit[26,27] (Thermo Scientific, 88229).

Biotinylated NET-DNA pull-down

DNA pull-down was performed as previously described[26,28]. In brief, cell membrane protein was incubated with 500 ng biotinylated NET-DNA in 400 µl IP lysis buffer (87787, Thermo Scientific) at room temperature for 1 h. The protein–DNA complex was then incubated with 50 µl streptavidin-agarose beads at room temperature for another 1 h. The beads were then washed 3 times with IP lysis buffer and separated through gradient gel electrophoresis followed by silver staining or western blotting. The specific band was analysed by liquid chromatography coupled to mass spectrometry.

NETs–CCDC25 binding assay

NETs (DNA–protein complex; 200 µg) were incubated with DNA-binding Dynabeads (4 mg ml^{-1}, Thermo Scientific, 37002D) at room temperature for 1 h and treated with PBS or proteinase K (0.5 µg ml^{-1}, Thermo Scientific) for 4 h at 56 °C. After centrifugation and extensive washing, the Dynabeads–NETs complex was incubated with 0.5 µg recombinant His–CCDC25 protein at room temperature for 1 h and the beads were washed four times with IP lysis buffer. Bound protein was eluted with 1X loading buffer by boiling for 5 min and resolved by 10% SDS–PAGE followed by immunoblotting with anti-His antibody.

In some experiments, 200 µg of NETs (DNA–protein complex) was incubated with the Biotin-XX sulfosuccinimidyl ester (Thermo Scientific, F-20650) on ice for 30 min and incubated with streptavidin-microbeads at room temperature for 1 h. After centrifugation and extensive washing, the Dynabeads–NETs complex was treated with PBS or DNase I (0.25 mg ml^{-1}, Roche, 11284932001) for 1 h at 37 °C and incubated with 0.5 µg recombinant His–CCDC25 protein. The beads were washed four times with IP lysis buffer. Bound protein was eluted with 1X loading buffer by boiling for 5 min, resolved by 10% SDS–PAGE and immunoblotted with anti-His antibody.

Electrophoretic mobility shift assay

The EMSA assay was performed according to the manufacturer' instructions using the LightShift Chemiluminescent EMSA Kit (Thermo Scientific, 20148). Biotinylated NET-DNA (1 ng) and 2 μg of isolated membrane protein or the indicated concentrations of recombinant protein His–CCDC25 were incubated in the EMSA binding buffer (Thermo Scientific). Specifically, a 20-fold excess of the unbiotinylated NET-DNA as competitor and 1:10 dilution of the anti-CCDC25 or IgG antibody were added and incubated for 40 min at room temperature. Thereafter, biotinylated NET-DNA was added to the mixture and incubated for 20 min at room temperature. For the competition EMSA, the corresponding unlabelled 8-OHdG-enriched DNA or non-8-OHdG-enriched DNA was used in addition to the biotinylated 8-OHdG DNA and His–CCDC25. The samples were applied to a 6% PAGE gel in 0.5 × TBE (Tris-borate-EDTA) buffer for 1.5 h at 100 V. The resolved reactions on the gel were transferred to a Nylon membrane for 1 h at 380 mA and the protein–DNA-binding complex was crosslinked to the membrane. The membrane was incubated with blocking solution at room temperature for 30 min to block non-specific binding and incubated with stabilized streptavidin-horseradish peroxidase at room temperature for 30 min. The membrane was then washed 4 times with wash solution and once with substrate equilibration buffer, and the presence of a band shift was assessed using chemiluminescent substrate working solution[29].

Bio-layer interferometry

Bio-layer interferometry assays were performed using an Octet RED96 instrument (ForteBio) at 25 °C in buffer containing 50 mM NaCl, 2 mM dithiothreitol and 20 mM Tris (pH 6.8). Anti-His biosensors (ForteBio) were pre-equilibrated in buffer for 10 min before each assay. Optimal sensor loading was achieved using 100 nM DNA or CCDC25 antibody and a loading period of 6 min. Dissociation constants (K_d) were determined from the binding data obtained with at least three concentrations of recombinant CCDC25. The data were analysed with the global fitting algorithm included in the Octet Data Analysis software (ForteBio)[18].

In vitro NET-DNA pull-down assay

HeLa cells transfected with the indicated plasmids were lysed in IP lysis buffer and centrifuged to obtain the supernatants. Supernatants were precipitated with anti-His antibody. The precipitated proteins were incubated with biotinylated NET-DNA for 1 h and additional 50 μl streptavidin beads for another 1 h at room temperature. The complex was washed 4 times with IP lysis buffer. The proteins were eluted with 1X loading buffer by boiling for 5 min and resolved by 10% SDS–PAGE followed by immunoblotting with anti-His antibody[30].

Heterologous DNA sequences

Three single-stranded sDNA sense strands containing a biotin label at the 3′ end (Oligo 1 (ref. [31]): CGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGT; Oligo 2 (ref. [28]): TACAGATCTACTAGTGATCTATGACTGATCTGTACATGATCTACATACAGATCTACTAGTGATCTATGACTGATCTGTACATGATCTACA; Oligo 3 (ref. [32]): GGGCTACCGTCAAGTAAGATGCAGATACGGAACACAGCTGGCACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACCCTCGGGATCT) were annealed to ssDNA antisense to create double-stranded DNA[28].

Oxidative DNA pull-down

Three heterologous double-stranded DNA oligonucleotides (Oligo 1: 5′-CGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGT-3′; Oligo 2: 5′-TACAGATCTACTAGTGATCTATGACTGATCTGTACATGATCTACATACAGATCTACTAGTGATCTATGACTGATCTGTACATGATCTACA-3′; Oligo 3: 5′-GGGCTACCGTCAAGTAAGATGCAGATACGGAACACAGCTGGCACAGTGGTAGTACTCCACTGTCTGGCTGTACAAAAACCCTCGGGATCT-3′) were dissolved in sterile H_2O at a concentration of 20 ng μl^{-1} and were irradiated with or without UV-C light at 250 mJ cm^{-2} using a UVP HL-2000 HybriLinker hybridization oven. The relative 8-OHdG content in the DNA was quantified with the 8-OHdG ELISA Kit (E-EL-0028c, Elabscience). The oxidative DNA pull-down was performed as follows: 0.5 μg recombinant His–CCDC25 protein was incubated with 200 ng oxidative or unmodified DNA in 400 μl IP lysis buffer at room temperature for 1 h. The protein–DNA complex was then incubated with 50 μl streptavidin-agarose beads at room temperature for another 1 h. The beads were then washed 3 times with IP lysis buffer and separated by gradient gel electrophoresis followed by immunoblotting with anti-His antibody.

Immunoprecipitation

HeLa cells were transfected with the indicated plasmids and were lysed in IP lysis buffer with protease inhibitor cocktail (78446, Thermo Scientific). Lysates were incubated with the indicated antibodies at room temperature for 1 h and an additional 50 μl of Dynabeads Protein A (Thermo Scientific, 10001D) for another 1 h at room temperature. The protein complex was washed 4 times with the IP lysis buffer, eluted with 1X loading buffer by boiling for 5 min and resolved by 10% SDS–PAGE followed by immunoblotting with the indicated antibodies.

Mammalian expression vectors

A cDNA encoding full-length human CCDC25, the N terminus of human CCDC25 (amino acids 1–168) and the C terminus of human CCDC25 (amino acids 41–208) were subcloned into the pcDNA4.1 vector with a C-terminal 6× His tag. To construct CCDC25 mutants, the five positively charged amino acid residues of human CCDC25 at positions 21–25 or 16–20 were substituted with alanine using QuikChange Lightning site-directed mutagenesis kit[33](Agilent Technologies). To evaluate the localization of CCDC25, we cloned full-length human CCDC25 constructs into pcDNA4.1 plasmid with an N-terminal Flag tag or a C-terminal Flag tag. For stably ectopic CCDC25 expression, the cDNA encoding full-length and mutant human CCDC25 were subcloned into the pLvx-IRES-Neo vector with a C-terminal Flag tag.

Membrane fractionation experiment

MDA-MB-231 cells (4×10^7) were collected and washed 3 times with cold PBS buffer and lysed in lysis buffer[20] (50 mM Tris/HCl pH 7.5, 0.3 M sorbitol,1 mM EDTA, 100 mM NaCl, 1 mM phenylmethylsulphonyl fluoride, 2 mM benzamidine and complete protease inhibitors cocktail) for 30 min. The total lysate was centrifuged at 20,000*g* for 40 min to obtain the soluble and pellet fractions. For membrane fractionation experiments, the pellet fraction was resuspended in lysis buffer containing 3 M NaCl, 5 M urea, 0.2 M Na_2CO_3 and 0.1% Triton X-100. After incubation for 30 min on ice, the samples were centrifuged at 20,000*g* for 40 min to separate the soluble and pellet fractions.

Flow cytometry

HeLa cells (5×10^5) were transfected with 2 μg of the plasmids encoding full-length human CCDC25 tagged with Flag on the N terminus or C terminus using Lipofectamine3000 according to the manufacturer's instructions (Invitrogen). Twenty-four hours after transfection, cells were collected, resuspended in PBS containing 1% FBS and stained with fluorescent-conjugated antibody against Flag-tag (1:100, BioLegend, 637309) for 30 min at 4 °C. For the cell permeabilization, Fixation and Permeabilization kit (88-8824-00, eBioscience) was used according to the manufacturer's instructions. All cells were analysed using a BD Accuri C6 Flow cytometer.

Immunohistochemistry

The paraffin-embedded samples were sectioned at 4-μm thickness. Antigen retrieval was performed using target retrieval solution, pH 9.0 (DAKO) using a pressure cooker for 15–20 min to remove the aldehyde links formed during the initial fixation of tissues. The non-specific binding was blocked with 5% BSA for 25 min at room temperature, the tissues were incubated with antibodies against CCDC25 (1:50, 21209-1-AP,ProteinTech) overnight at 4 °C and the immunodetection was performed using DAB (DAKO) according to the manufacturer's instructions.

Three-dimensional coculture

Cells were seeded within the 8-well Lab-Tek chambers. In brief, 40 μl Matrigel was added to the chambers and spread evenly. After the Matrigel solidified, 1,000 cells in 400 μl of assay medium containing 2% Matrigel with or without 5 μg ml^{-1} NET-DNA were added on the top of the solidified Matrigel and cultured for 4 days.

Adherence assay

Adherence of breast cancer cells to fibronectin or NETs was evaluated using 96-well plates that were pre-coated with 10 μg ml^{-1} fibronectin (Roche) or condition media containing 10 μg ml^{-1} NETs overnight at 37 °C. Breast cancer cells (5×10^4 per 100 μl) pretreated with 25 μg ml^{-1} mitomysin C (Sigma, M4287) for 2 h were then suspended in DMEM serum-free medium and were allowed to adhere to the plate bottom for 15 min (MDA-MB-231) or 60 min (MCF-7) at 37 °C. After removing the non-adherent cells by gently washing with PBS three times, the adhered cells were fixed in 4% paraformaldehyde for 20 min at room temperature and stained with crystal violet overnight at 4 °C. Cell adherence was counted as cells per field of view under phage-contrast microscopy.

Boyden chamber assay

Migration of breast cancer cells was examined using 24-well Boyden chambers (Corning) with 8 μm-inserts coated with fibronectin (Roche). Breast cancer cells (10^5 cells per well) pretreated with 25 μg ml^{-1} mitomysin C (Sigma. M4287) for 2 h were plated on the inserts and cultured at 37 °C in the upper chambers. After 4 h (MDA-MB-231 cells), the migrated cells that crossed the inserts were stained with crystal violet (0.005%, Sigma), and were counted as cells per field of view under phase-contrast microscopy.

Chemotaxis experiments

The μ-Slide Chemotaxis chamber (Ibidi) was used for quantitative chemotaxis experiments. The cancer cells were unlabelled or were labelled with carboxyfluorescein succinimidyl ester (CFSE) (0.5 μM, Thermo Scientific, C34554), Hoechst 33342, (5 μg ml^{-1}, Thermo Scientific, H21492) and CellTracker Orange CMTMR Dye (5 μM, Thermo Scientific, C2927) at 37 °C for 20 min. Cells were pipetted into the seeding chamber and the chambers were filled with PBS with or without 5 μg ml^{-1} NETs. The cells were allowed to settle for 1 h before examination by phase-contrast microscopy. Micro-images were captured every 20–30 min for 8–20 h. The cell migration tracks were analysed with ImageJ using a manual racking plugin and the chemotaxis and migration tool from Ibidi[34].

CRISPR-mediated gene knockout

The sequences targeting *CCDC25* or *ILK* were *CCDC25* gRNA1 (5′-GTCTGGGAACCGCTCGACTT-3′) and *CCDC25* gRNA2 (5′-GAATGCTATTGGCCTTCACA-3′); *ILK* gRNA1 (5′-CCACAGCAGAGCGGCCCTCT-3′) and *ILK* gRNA2 (5′-TCTCAACCGTATTCCATACA-3′). The Cas9 lentivirus and gRNA1/2 lentivirus were purchased from GenePharma and transduced to MDA-MB-231 tumour cells. The transduced cells were selected with 2.5 μg ml^{-1} puromycin for 2 weeks to obtain the *CCDC25* knockout MDA-MB-231 tumour cells. To rescue the expression of CCDC25 in *CCDC25*-knockout tumour cells[35], *CCDC25*-knockout cells were transduced with pLvx-IRES-Neo-CCDC25 (wild-type or with the indicated mutation) and selected in 800 μg ml^{-1} G418 for 2 weeks.

shRNA-mediated silencing

Tumour cells (2×10^5) were plated in 6-well plates and transduced with lentiviral particles (multiplicity of infection (MOI) of 5) with 8 μg ml^{-1} polybrene overnight at 37 °C. The cell supernatant was then removed and tumour cells were cultured in 2 ml of DMEM with 10% fetal bovine serum for 36 h. The transduced cells were selected with 2.5 μg ml^{-1} puromycin for two weeks. The shRNA target sequences are listed as follows: human *PARVB* (sh1), 5′-GGTGCTGGAAGCAGTACATGA-3′; human *PARVB* (sh2), 5′-GCATGTAACGGTGCAGGTGGT-3′.

Production of the polyclonal CCDC25 antibody

His-fused construct comprising the human CCDC25 protein was cloned into pET32a expression vector (EMD Millipore). The protein was expressed in *Escherichia coli* BL21(DE3) and purified by Ni-NTA agarose (88229, Thermo Scientific) according to the manufacturer's instructions. The protein was then immunized into rabbits and the antiserum was affinity-purified on antigen-coupled CNBr-activated agarose[36].

Application of anti-CCDC25 antibody in vivo

Neutralizing antibody against CCDC25 (1 mg kg^{-1})[37] was administrated concomitantly via tail vein every day when 1×10^6 MDA-MB-231-luc cells were intrasplenically injected into the spleen of NOD/SCID mice. Thirty days after tumour injection, mice were euthanized and the metastatic liver burden was measured using an IVIS Lumina Imaging System (Xenogen).

Immunoblotting

Protein was extracted from the cells with RIPA buffer and resolved on SDS–PAGE gels, then transferred to polyvinylidene difluoride membranes. The primary antibodies against CCDC25 (1:1,000, ProteinTech, 21209-1-AP), ILK (1:1,000, CST, 3862), HA-tag (1:1,000, Abcam, ab9110), His-tag (1:1,000, Thermo Scientific, MA1-21315), β-parvin (1:1,000, ProteinTech, 14463-1-AP), α-parvin (1:1,000, CST, 4026), Pinch (1:1,000, CST, 11890), H3cit (1:1,000, Abcam, ab5103), CXCR4 (1:1,000, ProteinTech, 60042-1-Ig), ATP5b (1:1,000, ProteinTech, 17247-1-AP) and GAPDH (1:10,000, ProteinTech, HRP-60004) were used. Peroxidase-conjugated secondary antibody (CST) was used and the antigen–antibody reaction was visualized using an enhanced chemiluminescence assay (ECL, Thermo). To analyse the activity of the Rho-family GTPases (CDC42 and RAC1), cells were stimulated with condition media containing 5 μg ml^{-1} NETs and the purification of active, GTP-bound protein, as well as the subsequent immunoblotting, was performed using RAC1/CDC42 Assay Reagent (PAK1 PBD, agarose; Millipore) according to the manufacturer's protocol[38].

Statistical analysis

All the statistical analyses were performed using GraphPad Prism 7 software, and error bars indicate s.e.m. or s.d. The number of independent experiments, the number of events and information about the statistical details and methods are indicated in the relevant figure legends.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Source data are provided with this paper. All other data are available from the corresponding author upon reasonable request.

21. Yipp, B. G. et al. Infection-induced NETosis is a dynamic process involving neutrophil multitasking in vivo. *Nat. Med*. **18**, 1386–1393 (2012).
22. Nie, M. et al. Neutrophil extracellular traps induced by IL8 promote diffuse large B-cell lymphoma progression via the TLR9 signaling. *Clin. Cancer Res*. **25**, 1867–1879 (2018).
23. Su, S. et al. CD10$^+$GPR77$^+$ cancer-associated fibroblasts promote cancer formation and chemoresistance by sustaining cancer stemness. *Cell* **172**, 841–856.e16 (2018).
24. Baer, C. et al. Suppression of microRNA activity amplifies IFN-γ-induced macrophage activation and promotes anti-tumour immunity. *Nat. Cell Biol*. **18**, 790–802 (2016).
25. Wang, C. Y. et al. ZIP8 is an iron and zinc transporter whose cell-surface expression is up-regulated by cellular iron loading. *J. Biol. Chem*. **287**, 34032–34043 (2012).
26. Roberts, T. L. et al. HIN-200 proteins regulate caspase activation in response to foreign cytoplasmic DNA. *Science* **323**, 1057–1060 (2009).
27. Zhang, Z. et al. The helicase DDX41 senses intracellular DNA mediated by the adaptor STING in dendritic cells. *Nat. Immunol*. **12**, 959–965 (2011); correction **13**, 196 (2012).
28. Abe, T. et al. STING recognition of cytoplasmic DNA instigates cellular defense. *Mol. Cell* **50**, 5–15 (2013).
29. Xu, R. H. et al. *NANOG* is a direct target of TGFβ/activin-mediated SMAD signaling in human ESCs. *Cell Stem Cell* **3**, 196–206 (2008).
30. Zhang, Z. et al. DDX1, DDX21, and DHX36 helicases form a complex with the adaptor molecule TRIF to sense dsRNA in dendritic cells. *Immunity* **34**, 866–878 (2011).
31. Jensen, R. B., Carreira, A. & Kowalczykowski, S. C. Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature* **467**, 678–683 (2010).
32. Lu, H., Pannicke, U., Schwarz, K. & Lieber, M. R. Length-dependent binding of human XLF to DNA and stimulation of XRCC4·DNA ligase IV activity. *J. Biol. Chem*. **282**, 11155–11162 (2007).
33. Yamamoto, Y. & Sakisaka, T. Molecular machinery for insertion of tail-anchored membrane proteins into the endoplasmic reticulum membrane in mammalian cells. *Mol. Cell* **48**, 387–397 (2012).
34. Hotz, M. & Nelson, W. J. Pumilio-dependent localization of mRNAs at the cell front coordinates multiple pathways required for chemotaxis. *Nat. Commun*. **8**, 1366 (2017).
35. Zhu, Y., Wang, G. Z., Cingöz, O. & Goff, S. P. NP220 mediates silencing of unintegrated retroviral DNA. *Nature* **564**, 278–282 (2018).
36. Li, M. et al. The histone modification reader ZCWPW1 is required for meiosis prophase I in male but not in female mice. *Sci. Adv*. **5**, eaax1101 (2019).
37. Tvorogov, D. et al. Effective suppression of vascular network formation by combination of antibodies blocking VEGFR ligand binding and receptor dimerization. *Cancer Cell* **18**, 630–640 (2010).
38. Shibue, T., Brooks, M. W. & Weinberg, R. A. An integrin-linked machinery of cytoskeletal regulation that enables experimental tumor initiation and metastatic colonization. *Cancer Cell* **24**, 481–498 (2013).

Acknowledgements This work was supported by grants from the National Key Research and Development Program of China (2016YFC1302300, 2017YFA0106300), the Natural Science Foundation of China (81621004, 81490750, 81720108029, 81930081, 91942309, 81672614, 81802643, 81902688, 81902699, 81972465, 81802645, 91940305), Guangdong Science and Technology Department (2017B030314026), Clinical Innovation Research Program of Guangzhou Regenerative Medicine and Health Guangdong Laboratory (2018GZR0201004), Guangzhou Science Technology and Innovation Commission (201803040015, 201508020008, 201508020249), Natural Science Foundation of Guangdong Province (2016A030306023, 2017A030313878, 2018A030310085, 2019A1515011485), Pearl River S&T Nova Program of Guangzhou (201710010083), Tip-top Scientific and Technical Innovative Youth Talents of Guangdong Special Support Program (2016TQ03R553) and the China Postdoctoral Science Foundation (BX20190396, 2019M663270, 2018M640868, 2019TQ0367, 2019M660228). This research is partly supported by Fountain-Valley Life Sciences Fund of University of Chinese Academy of Sciences Education Foundation.

Author contributions L.Y., Q.L., S.S. and E.S. conceived the ideas, designed the experiments and wrote the manuscript. L.Y. performed most of the experiments and analysed the data. The bio-layer interferometry assays were reproduced independently by X.Z. The flow cytometry assays and three-dimensional coculture assays were performed by L.Y. and X.L. and were reproduced by J. Li. The adhesion assays and Boyden chamber assays were performed by L.Y. and B.Z. and reproduced by Y.X. Chemotaxis experiments were performed by J.C. and reproduced by X.C. and J. Liu. Immunoblotting assays were reproduced independently by H.L. Q.L. and E.S. provided samples from patients for clinical data analysis. D.H. and F.C. performed the NET quantification independently. All authors contributed to the revision of the manuscript.

Competing interests The authors declare no competing interests.

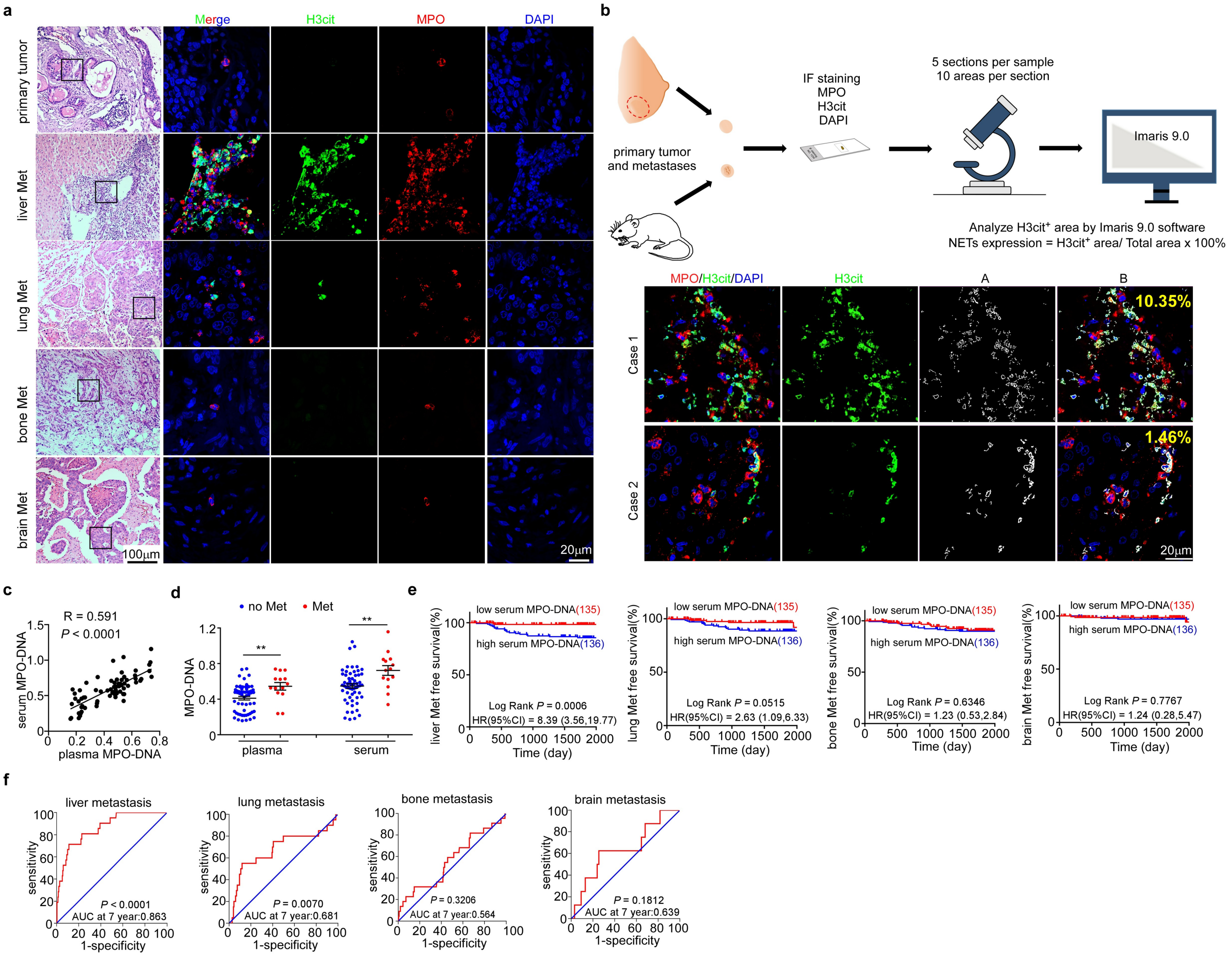
Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2394-6.

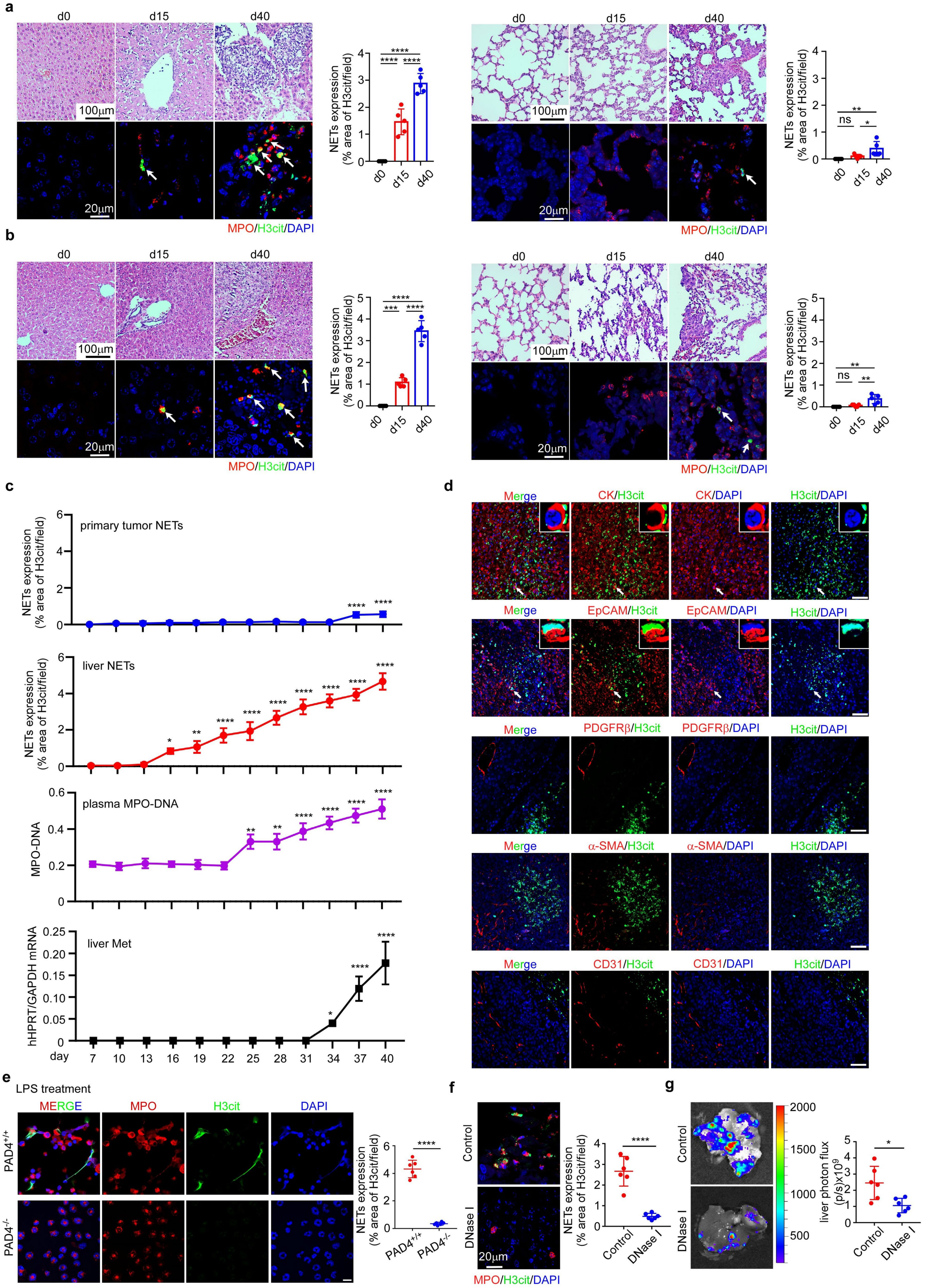
Correspondence and requests for materials should be addressed to S.S. or E.S.

Peer review information *Nature* thanks Ilaria Malanchi and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at http://www.nature.com/reprints.

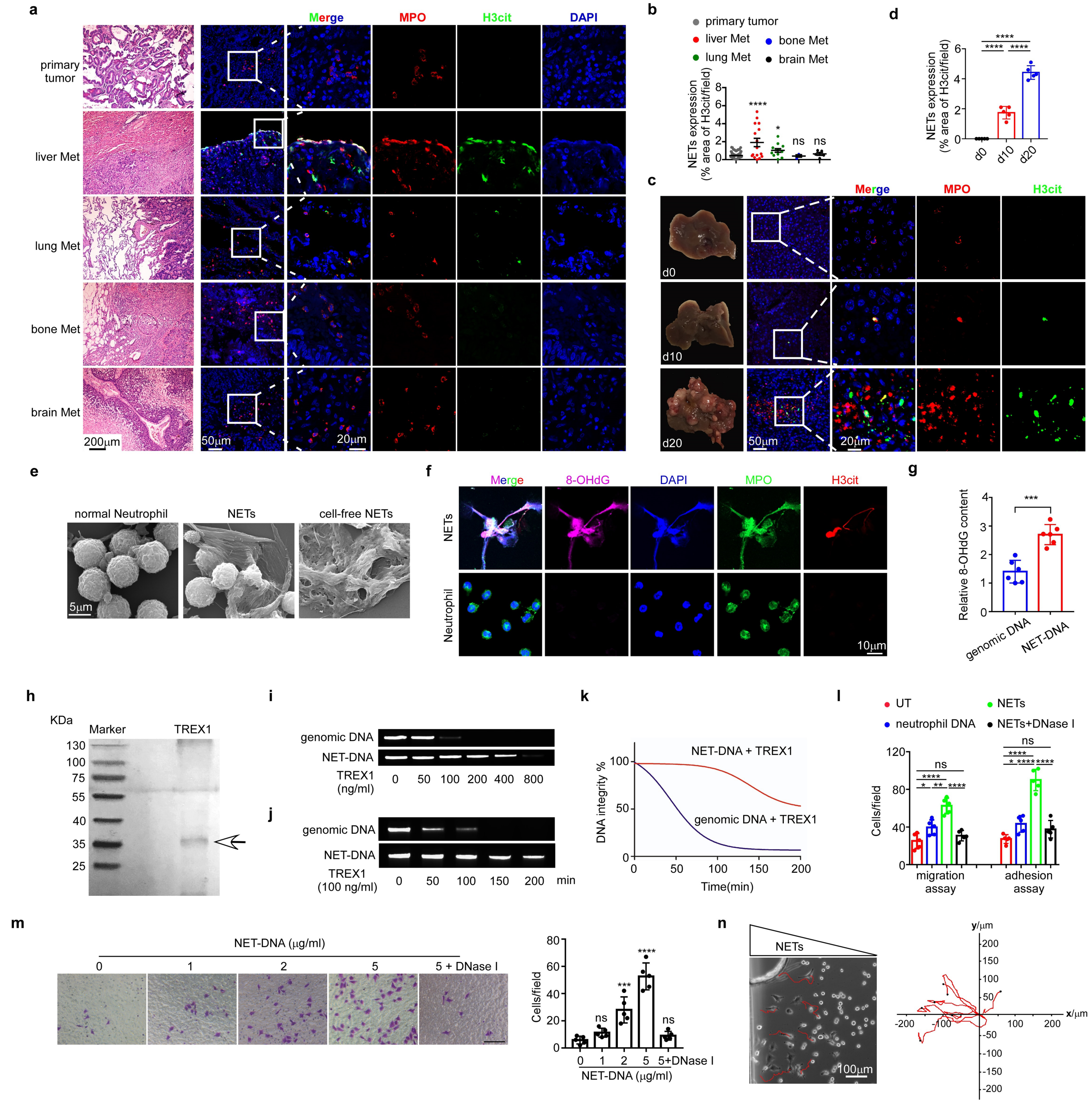


Extended Data Fig. 1 | NETs are predominantly presented in liver metastases of breast cancer. a, Representative images of haematoxylin and eosin (H&E) staining (first column) and immunofluorescence staining for myeloperoxidase (red), citrullinated histone H3 (green) and DAPI (blue) (subsequent columns) in human primary breast cancer (*n* = 461) and metastases (Met) in liver (*n* = 20), lung (*n* = 23), brain (*n* = 7) or bone (*n* = 33). **b**, NET quantification was performed by immunofluorescence staining using Imaris 9.0 Microscopy Image Analysis Software. The first column indicates MPO, H3cit and DAPI staining and the second column indicates H3cit staining in the same tissue section. Columns A and B show the results of analysis using the Imaris 9.0 Software. Column A indicates the H3cit-positive signal area, and column B shows the percentage of H3cit areas in the whole section. **c**, Correlation between serum MPO–DNA and plasma MPO–DNA levels in breast cancer samples (*n* = 72, the Pearson's correlation coefficient *R* value and the *P* value are shown). **d**, Plasma and serum levels of MPO–DNA in patients with breast cancer with (*n* = 14) or without (*n* = 58) distant organ metastases. Data are mean ± s.e.m., ***P* = 0.0052 (plasma) and 0.0035 (serum), calculated using two-tailed Student's *t*-test. **e**, Kaplan–Meier survival curves for patients with breast cancer with low (*n* = 135) and high (*n* = 136) serum MPO–DNA levels. The significance was assessed using a two-sided long-rank test. **f**, Receiver operator characteristic (ROC) curves to predict liver, lung, bone or brain metastases from serum MPO–DNA levels. *n* = 271, AUC, area under curve.



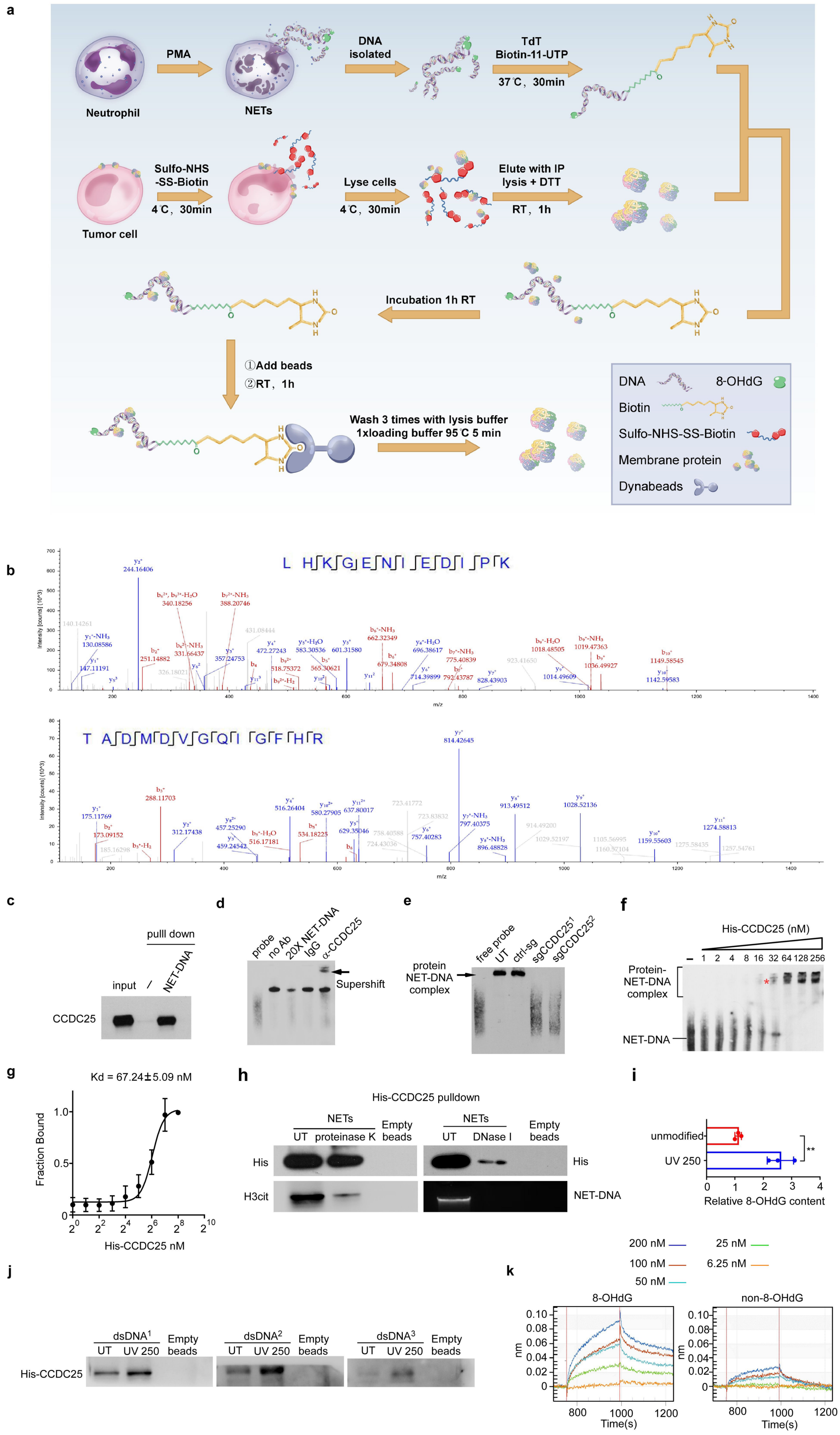
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | NETs promote liver metastases. a, **b**, Mouse 4T1 (**a**) and human MDA-MB-231 (**b**) breast cancer cells were injected into mammary fat pads of BALB/c mice (**a**) and NOD/SCID mice (**b**), respectively. At various time points (0, 15 and 40 days) after tumour inoculation, the mice were killed and examined for NET infiltration and tumour metastases in the liver and lungs. Representative images of H&E staining and immunofluorescence staining for H3cit and MPO to denote NET infiltration in the liver (left) and the lung (right) are shown, white arrows indicate NETs. $n = 5$ per group. Data are mean ± s.d., two-sided one-way ANOVA with Tukey test; $^{****}P < 0.0001$; for **a**, ns = 0.5985, $^{*}P = 0.0422$, $^{**}P = 0.0072$; for **b**, $^{***}P = 0.0003$, ns = 0.7300, $^{**}P = 0.0016$ (day 0) and 0.0063 (day 15) compared with day 40. **c**, MDA-MB-231 breast cancer cells were injected into mammary fat pads of NOD/SCID mice, and the tumour tissues, liver tissues and plasma were collected at different time points after tumour inoculation. The dynamics of NET expression in the primary tumours and the liver, the plasma MPO–DNA levels and the expression of liver *HPRT1* mRNA relative to mouse *Gapdh* expression were shown ($n = 3$ mice per group). Data are mean ± s.d., two-sided one-way ANOVA with Tukey test. $^{****}P < 0.0001$; for liver NETs group: $^{*}P = 0.0478$, $^{**}P = 0.0064$; for plasma MPO–DNA group: $^{**}P = 0.0014$; for liver Met group: $^{*}P = 0.0454$, compared with d7 group. **d**, Representative images of immunofluorescence co-staining for H3cit with CK and epithelial cell adhesion molecule (EpCAM) for tumour cells, platelet-derived growth factor receptor beta (PDGFRβ) for pericytes, α-smooth muscle actin (α-SMA) for stromal cells and CD31 for endothelial cells in the metastatic liver tissues of NOD/SCID mice intrasplenically injected with MDA-MB-231 cells. White arrows indicate the areas that are shown in the higher-magnification images in the top-right corners. Scale bars, 50 µm. $n = 3$ biologically independent experiments. **e**, Representative images and quantification of NETs in the LPS-induced neutrophils isolated from *PAD4*$^{+/+}$ and *PAD4*$^{-/-}$ mice. Scale bars, 10 µm. $n = 6$ biologically independent animals. Data are mean ± s.d. $^{****}P < 0.0001$, assessed using a two-tailed Student's *t*-test. **f**, **g**, Representative images (left) and quantification (right) of liver NET formation (**f**) and liver metastases (**g**) of luciferase-E0771 tumour cells injected into the spleens (1×10^6 cells per mouse of wild-type and DNase I-treated female C57BL/6 mice ($n = 6$ mice per group). Data are mean ± s.d. $^{*}P = 0.0122$, $^{****}P < 0.0001$ assessed using a two-tailed Student's *t*-test.



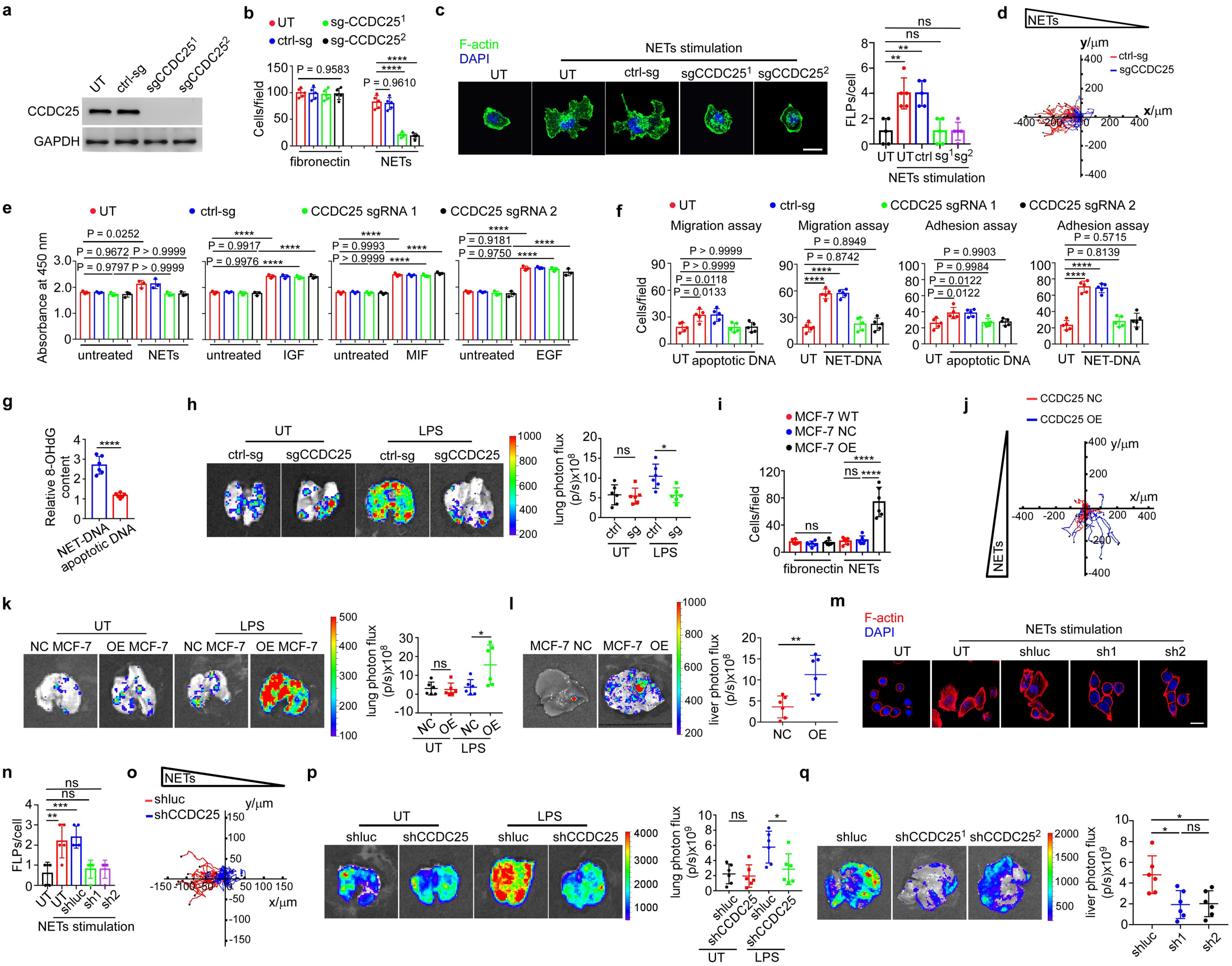
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | 8-OHdG-enriched NETs are predominantly detected in liver metastases of colon cancer. a, Representative images of H&E (first column) and immunofluorescence (subsequent columns) staining for MPO (red), H3cit (green) and DAPI (blue) in human primary colon cancer and metastases (Met) in the liver, lung, bone or brain. **b**, NETs infiltrated in primary colon cancer tissues (n = 130) and in liver (n = 16), lung (n = 12), bone (n = 3) and brain (n = 5) metastases. Data are mean ± s.e.m., two-sided one-way ANOVA with Tukey test, $^{****}P < 0.0001$, $^{*}P = 0.0359$, ns > 0.9999 (bone met.) and = 0.9710 (brain met.) compared with primary tumour. Met, metastases. **c**, **d**, Representative images of confocal microscopy (**c**) and quantification (**d**) of NETosis, denoted by H3cit and MPO immunofluorescence staining in the liver tissues at various time points (0, 10 and 20 days) following intrasplenic injection of HCT116 colon cancer cells. n = 5 mice per time point. Data are mean ± s.d., two-sided one-way ANOVA with Tukey test, $^{****}P < 0.0001$. **e**, Representative scanning electron microscopy images of normal or PMA-stimulated neutrophils (NETs) and cell-free NETs isolated from PMA-stimulated neutrophils (cell-free NETs). **f**, Representative images of 8-OHdG staining in the NETs (top) produced by PMA-stimulated neutrophils or normal neutrophils (bottom). **g**, 8-OHdG levels in the genomic and NET-DNA of human neutrophils, determined by 8-OHdG ELISA assays (n = 6 biologically independent samples). Data are mean ± s.d. $^{***}P = 0.0001$ calculated using a two-tailed Student's t-test. **h**, Silver staining of His–TREX1 expressed and purified from *Escherichia coli*. **i**, **j**, Agarose gel analysis of the genomic DNA and NET-DNA from neutrophils incubated with increasing concentrations of recombinant TREX1 protein (**i**) or with 100 ng ml^{-1} of recombinant TREX1 protein for increasing time periods (**j**). **k**, Dynamics of the levels of genomic DNA and NET-DNA treated with recombinant TREX1 at 100 ng ml^{-1}, quantified from the agarose gels in **j**. **l**, Adhesion and migration assays for MDA-MB-231 cells stimulated with 5 μg ml^{-1} neutrophil DNA or 5 μg ml^{-1} NETs in the presence or the absence of DNase I. n = 5 biologically independent experiments. Data are mean ± s.d., two-sided one-way ANOVA with Tukey test, $^{****}P < 0.0001$; for migration assays: $^{*}P = 0.0390$, $^{**}P = 0.0011$ and ns = 0.6578; for adhesion assay: $^{*}P = 0.0469$ and ns = 0.2841. **m**, Migration assays for MDA-MB-231 cells in Boyden chambers. NET-DNA at increasing concentrations (0–5 μg ml^{-1}) or pretreated with DNase I was added to the culture media in the lower chambers. n = 5 biologically independent experiments, Data are mean ± s.d. two-sided one-way ANOVA with Tukey test, $^{****}P < 0.0001$, $^{***}P = 0.0003$, ns = 0.6978 (1 μg ml^{-1}) and 0.9372 (5 μg ml^{-1} + DNase I) compared with the untreated cells. Scale bars, 100 μm. **n**, MDA-MB-231 cells were randomly attached to the seeding chamber in PBS. The media in the left chamber was replaced with media containing 5 μg ml^{-1} NET-DNA. Tracks of individual cells are shown as coloured lines (left). The spider plot (right) demonstrates tracks of the bulk cells. Data in **e**, **f**, **h**–**k**, **n** were representative of three biologically independent experiments.



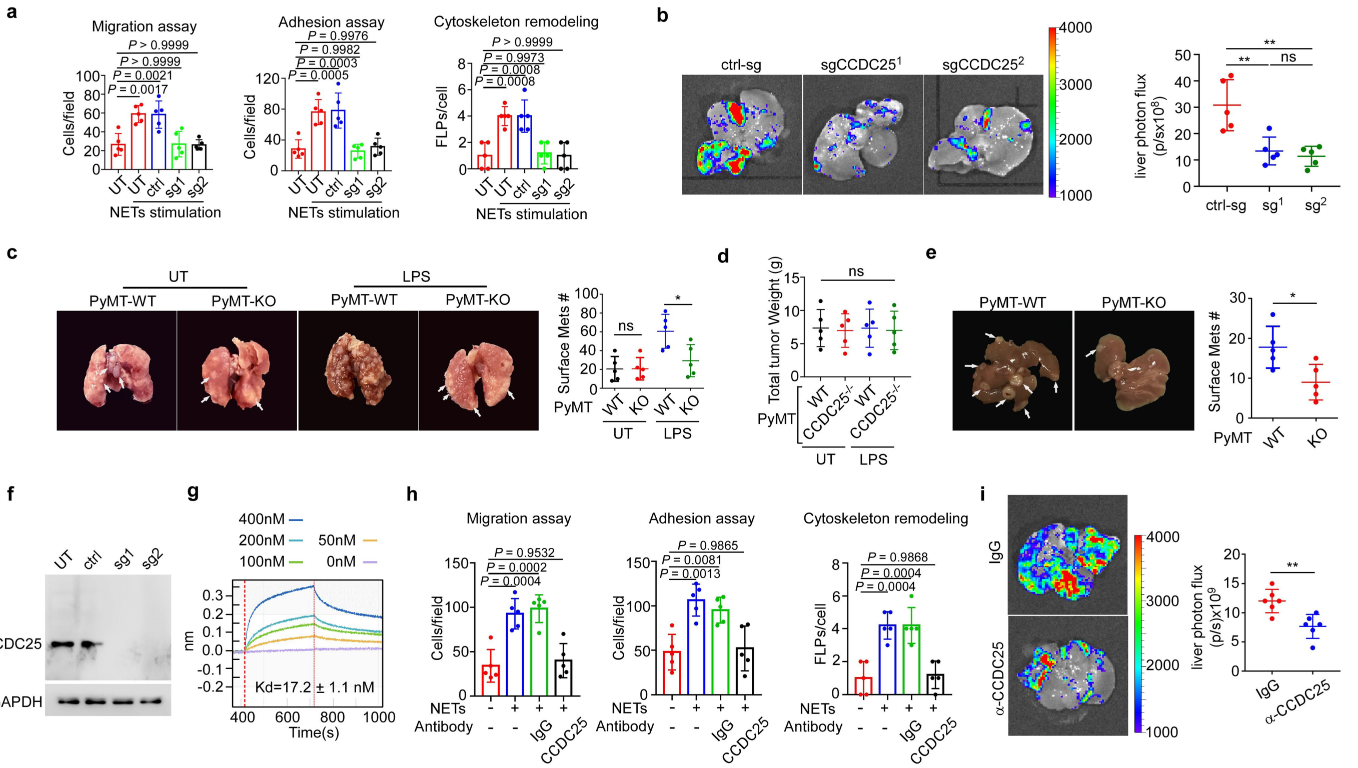
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | CCDC25 binds to NET-DNA. a, Schematic of NET-DNA pull-down assays. **b**, Mass spectrometry analysis identified CCDC25 as the cytoplasmic membrane protein from MDA-MB-231 cells pulled down by the biotinylated NET-DNA. **c**, Immunoblotting of the membrane proteins of MDA-MB-231 cells pulled down by biotinylated NET-DNA and detected by an anti-CCDC25 antibody. **d**, EMSA demonstrated NET-DNA binding to CCDC25 super-shifted by an anti-CCDC25 antibody. Membrane proteins of MDA-MB-231 cells and the biotinylated NET-DNA were incubated with or without the antibody against CCDC25, IgG (negative control), or 20-fold excess of unbiotinylated NET-DNA. **e**, The binding of NET-DNA to the membrane proteins of MDA-MB-231 cells transduced with a control sgRNA or two CCDC25-sgRNAs was evaluated by EMSA. **f**, EMSA reveals the interaction of biotinylated NET-DNA with increasing concentrations of CCDC25. The protein–DNA complex is denoted by a red asterisk. **g**, Binding kinetics of CCDC25 and NET-DNA generated from the above EMSA assays in **f**. $n = 3$. Data are mean ± s.d. **h**, Purified NETs were coupled to magnetic beads, treated with Proteinase K (left) and DNase I (right), and incubated with His–CCDC25. The interaction of NETs and CCDC25 was evaluated by the precipitation of NETs–beads and blotted with anti-His antibody. His–CCDC25 mixed with beads without DNA served as a negative control (empty beads). The digestion efficiency of the protein and DNA components of NETs by Proteinase K and DNase I was confirmed by immunoblotting for H3cit and agarose gel analysis for DNA. **i**, **j**, Three different biotinylated heterologous 90-bp DNA duplexes with random sequences were either irradiated with UV-C light or were not irradiated. **i**, The relative 8-OHdG content in the DNA was determined by ELISA. $n = 3$. Data are mean ± s.d., $^{**}P = 0.0059$ as determined by a two-tailed Student's *t*-test. **j**, DNA pull-down assay for His–CCDC25. The resultant CCDC25 was detected by anti-His western blot analysis. His–CCDC25 mixed with beads without DNA served as a negative control (empty beads). **k**, Representative bio-layer interferometry showing CCDC25 binding to 8-OHdG-enriched DNA (left) and non-8-OHdG-enriched DNA (Right). The coloured lines show the data for five different concentrations of CCDC25 as indicated. Data in **b**–**f**, **h**–**k** are representative of three biologically independent experiments.

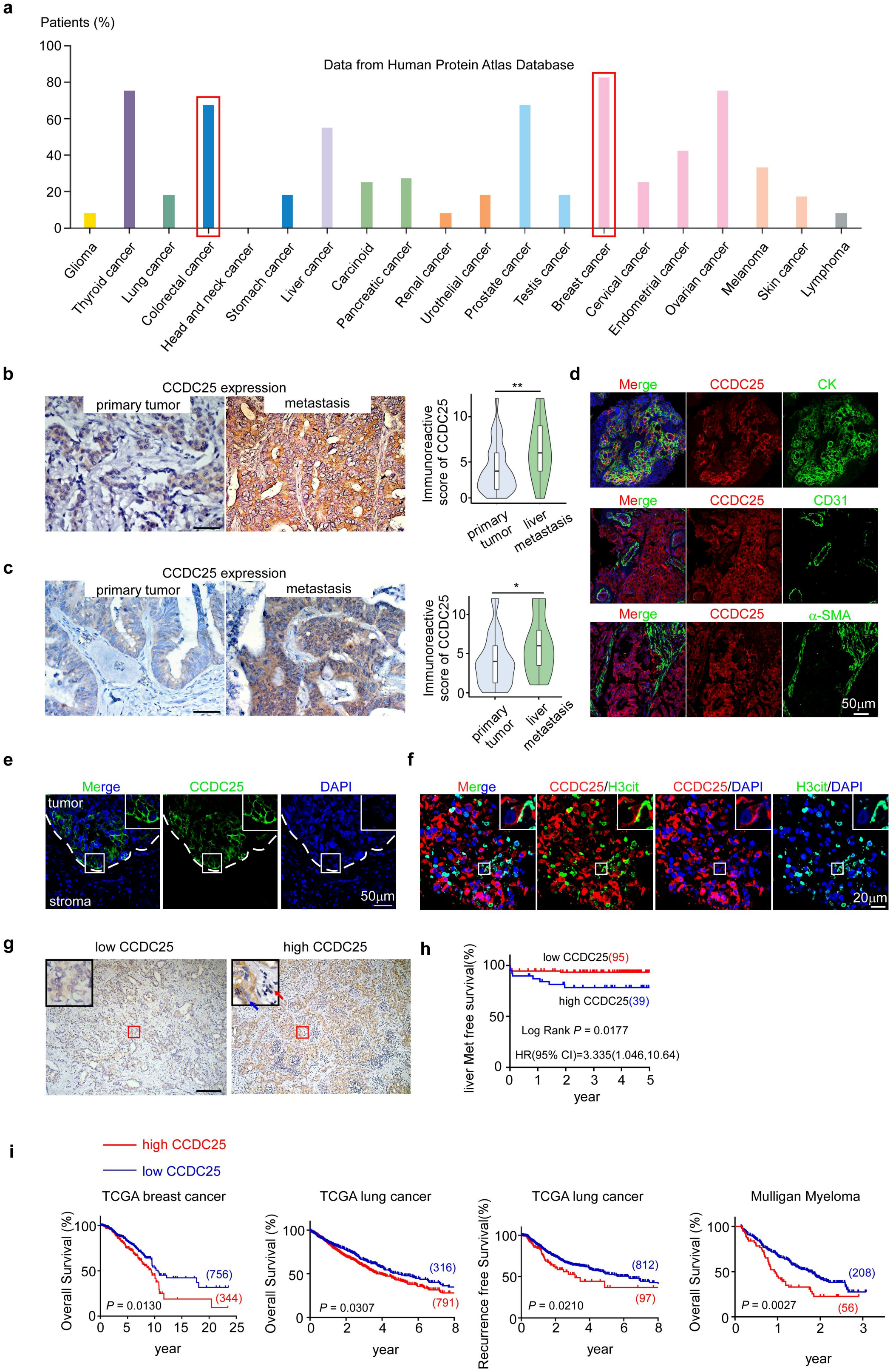


Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | NETs promote tumour metastasis via CCDC25. a, Western blots for CCDC25 expression in MDA-MB-231 cells that were untreated (UT), transfected with a control sgRNA (ctrl-sg) or with one of two *CCDC25* sgRNAs (sgCCDC25[1] and sgCCDC25[2]). **b**, MDA-MB-231 cells untreated (UT) or transfected with a control sgRNA or one of two *CCDC25* sgRNAs were added into culture plates coated with fibronectin or NETs, and cell adhesion was evaluated. $n = 5$ biologically independent experiments. Data are mean ± s.d., two-sided one-way ANOVA with Tukey test, $^{****}P < 0.0001$. **c**, MDA-MB-231 cells transduced with either a control sgRNA or one of two CCDC25-sgRNAs were treated with 5 μg ml^{-1} NET-DNA or left untreated (UT), and stained with phalloidin (F-actin, green) and DAPI (nuclei, blue). Scale bars, 20 μm. FLPs, filopodium-like protrusions. $n = 5$ biologically independent experiments. Data are mean ± s.d., two-sided one-way ANOVA with Tukey test, $^{**}P = 0.0011$, ns >0.9999. **d**, MDA-MB-231 cells were transduced with a control sgRNA or with CCDC25-sgRNA, and monitored for migration in a chemotaxis chamber with NET-DNA on the left. Red and blue lines demonstrate the migration tracks of control and CCDC25-knockout tumour cells, respectively. **e**, MDA-MB-231 cells transduced with a control sgRNA or with one of two sgRNAs for CCDC25 were treated with 5 μg ml^{-1} NET-DNA, 50 ng ml^{-1} IGF, 50 ng ml^{-1} MIF or 50 ng ml^{-1} EGF or were left untreated. Cell proliferation was assessed by the Cell Counting Kit-8 (CCK-8) assay. $n = 3$ biologically independent experiments. Data are mean ± s.d., two-sided one-way ANOVA with Tukey test. $^{****}P < 0.0001$. **f**, MDA-MB-231 cells transduced with a control sgRNA or with one of two sgRNAs for CCDC25 were treated with 5 μg ml^{-1} NET-DNA or 5 μg ml^{-1} apoptotic DNA or were left untreated, and cell migration and adhesion were evaluated. $n = 5$ biologically independent experiments. Data are mean ± s.d., two-sided one-way ANOVA with Tukey test. $^{****}P < 0.0001$. **g**, 8-OHdG levels in NET-DNA and apoptotic DNA, determined by 8-OHdG ELISA assays, $n = 6$ biologically independent experiments. Data are mean ± s.d. $^{****}P < 0.0001$ as calculated by a two-tailed Student's *t*-test. **h**, Luciferase-MDA-MB-231 cells transduced with a control or CCDC25-targeting sgRNA were intravenously injected into NOD/SCID mice that were pretreated with LPS or untreated; representative images and quantification of lung metastases in mice with the indicated treatments are shown ($n = 6$ mice per group). Two-sided one-way ANOVA with Tukey test, $^{*}P = 0.0132$, ns = 0.9958. **i**, MCF-7 cells that were untreated (MCF-7 WT), or transduced with negative control (MCF-7 NC) or with CCDC25-overexpression vectors (MCF-7 OE) were added into the culture plates coated with fibronectin or NETs, and cell adhesion was evaluated. $n = 6$ biologically independent experiment. Data are mean ± s.d. Two-sided one-way ANOVA with Tukey test. ns = 0.4764 (fibronectin group) and 0.9744 (NETs group), $^{****}P < 0.0001$. **j**, Migration tracks of the MCF-7 cells transduced with negative control (CCDC25 NC) or with CCDC25-overexpression vectors (CCDC25 OE) in a chemotaxis chamber containing culture media with 5 μg ml^{-1} NETs. Red and blue lines demonstrate the tracks of control and CCDC25-overexpressed tumour cells, respectively. **k**, Luciferase-MCF-7 cells with (OE MCF-7) or without (NC MCF-7) CCDC25 overexpression were intravenously injected into NOD/SCID mice that were pretreated with LPS or were untreated. Representative images and quantification of lung metastases in the mice with indicated treatments are shown ($n = 6$ mice per group). Data are mean ± s.d., two-sided one-way ANOVA with Tukey test. ns = 0.9989, $^{*}P = 0.0170$. **l**, Representative images and quantification of liver metastases in NOD/SCID mice that were intrasplenically injected with luciferase MCF-7 cells with (MCF-7 OE) or without (MCF-7 NC) CCDC25 overexpression, $n = 6$ mice per group). Data are mean ± s.d., $^{**}P = 0.0052$ as calculated using a two-tailed Student's *t*-test. **m**, **n**, HCT116 cells transduced with a control shRNA or one of two CCDC25-shRNAs were treated with 5 μg ml^{-1} NET-DNA or were left untreated, and were stained with phalloidin (F-actin, red) and DAPI (nuclei, blue). Scale bars, 10 μm. $n = 5$ biologically independent experiments. Data are mean ± s.d., two-sided one-way ANOVA with Tukey test. $^{**}P = 0.0026$, $^{***}P = 0.0008$, ns = 0.9817. **o**, HCT116 cells were transduced with a control shRNA (shluc) or with CCDC25-shRNA, and monitored for migration in a chemotaxis chamber with NET-DNA on the left. Red and blue lines show the migration tracks of control and CCDC25-knockdown tumour cells, respectively. **p**, Luciferase-HCT116 cells transduced with a control shRNA or with CCDC25-targeting shRNA were intravenously injected into NOD/SCID mice pretreated with LPS or untreated. Representative images and quantification for lung metastases in mice with the indicated treatments are shown ($n = 6$ mice per group). Data are mean ± s.d., two-sided one-way ANOVA with Tukey test. $^{*}P = 0.0435$, ns = 0.9922. **q**, NOD/SCID mice were intrasplenically injected with luciferase-HCT116 cells, which were transduced with a control or with one of two CCDC25-targeting shRNAs. Representative images and quantification of liver metastases with indicated treatments were shown ($n = 6$ mice per group). Data are mean ± s.d., two-sided one-way ANOVA with Tukey test. $^{*}P = 0.0117$ (sh[1] versus shluc) and 0.0142 (sh[2] versus shluc), ns = 0.9948. Data in **a**, **d**, **j**, **o** are representative of three independent experiments.

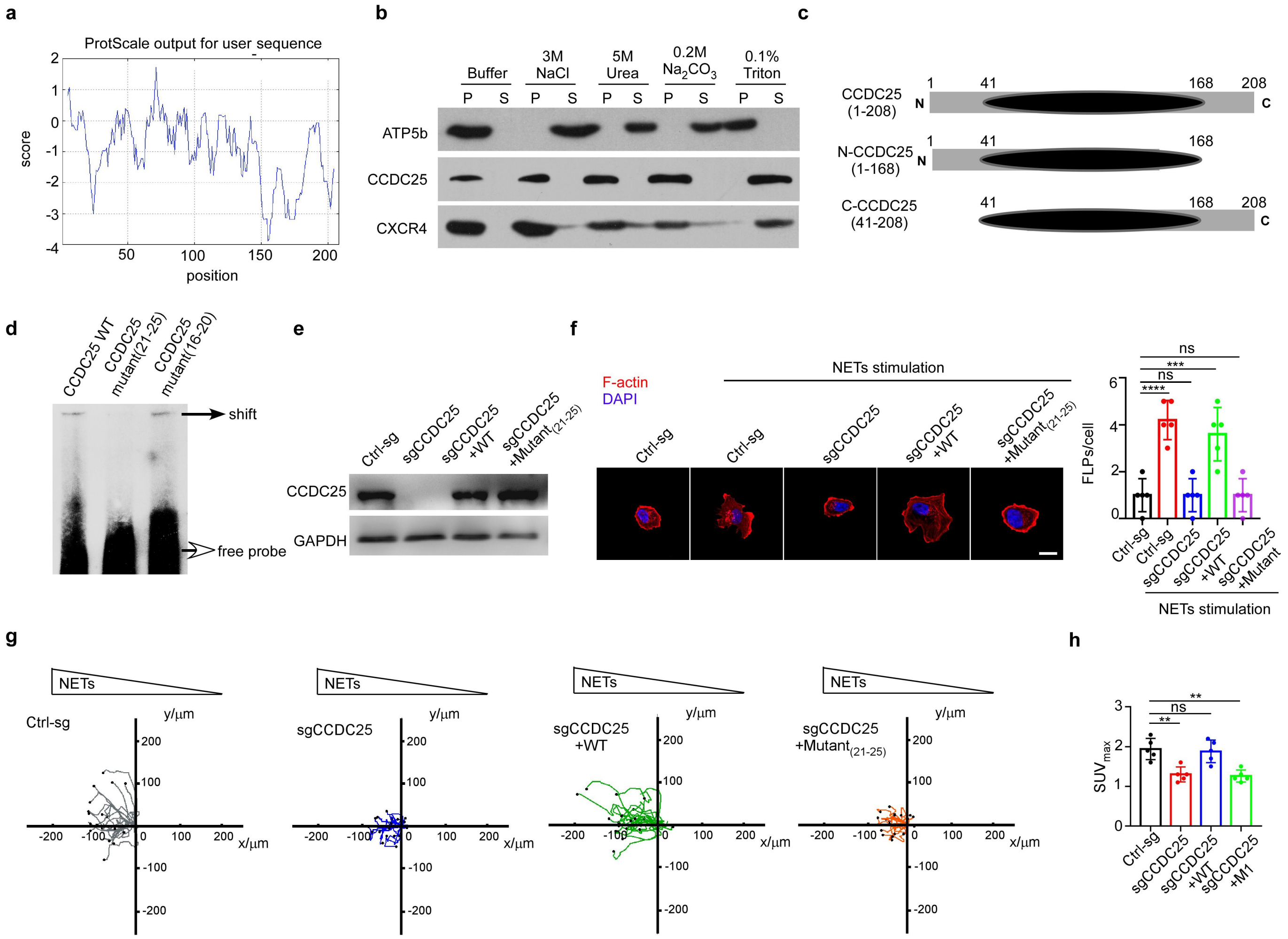


Extended Data Fig. 6 | NETs promote tumour metastasis via CCDC25. **a**, Primary breast cancer cells transduced with a control sgRNA or with one of two sgRNAs for CCDC25 were treated with 5 μg ml^{-1} NET-DNA or were untreated, and cell migration, adhesion and cytoskeleton remodelling were evaluated. $n = 5$ biologically independent experiments. Data are mean ± s.d. two-sided one-way ANOVA with Tukey test. **b**, NOD/SCID mice were intrasplenically injected with luciferase-primary breast cancer cells transduced with a control sgRNA or with one of two sgRNAs for CCDC25. Representative images (left) and quantification (right) of liver metastases after the indicated treatments are shown ($n = 5$ mice per group). Data are mean ± s.d., two-sided one-way ANOVA with Tukey test, $^{**}P = 0.0040$ (sg^1 versus ctrl-sg) and 0.0018 (sg^2 versus ctrl-sg), ns = 0.8865. **c**, Representative images (left) and quantification (right) of lung metastases in wild-type (WT) PyMT mice and in PyMT;CCDC25$^{-/-}$ (PyMT-KO) mice pretreated with LPS or untreated ($n = 5$ mice per group). Data are mean ± s.d., two-sided one-way ANOVA with Tukey test, $^{*}P = 0.0244$, ns >0.9999. **d**, Tumour burden in wild-type PyMT mice and PyMT;CCDC25$^{-/-}$ mice pretreated with LPS or untreated. $n = 5$ mice per group. Data are mean ± s.d., two-sided one-way ANOVA, ns = 0.9934. **e**, Representative images (left) and quantification (right) of liver metastases in C57BL/6 mice intrasplenically injected with tumour cells derived from the wild-type PyMT mice and PyMT-KO mice ($n = 5$ mice per group). Data are mean ± s.d., $^{*}P = 0.0215$ determined by a two-tailed Student's *t*-test. **f**, Validation of the polyclonal CCDC25 antibody. Indicated cell lysates from MDA-MB-231 cells transduced with a control sgRNA or with one of two sgRNAs for CCDC25 were subjected to western blot analysis probing with the polyclonal CCDC25 antibody. **g**. Representative bio-layer interferometry data of polyclonal antibody binding to recombinant protein CCDC25. The coloured lines show the data for five different concentrations of recombinant CCDC25 as indicated. **h**, Inhibitory effects of a polyclonal CCDC25 blocking antibody (5 μg ml^{-1}) on NET-induced migration, adhesion and cytoskeleton arrangement of MDA-MB-231 cells. $n = 5$ biologically independent experiments. Data are mean ± s.d., two-sided one-way ANOVA with Tukey test. **i**, Representative images and quantification of liver metastases of NOD/SCID mice intrasplenically injected with luciferase-MDA-MB-231 cells, which were treated with IgG as a control or with a CCDC25 antibody. $n = 6$ mice per group. Data are mean ± s.d., $^{**}P = 0.0042$ as calculated by a two-tailed Student's *t*-test. Data in **f**, **g** are representative of three independent experiments.

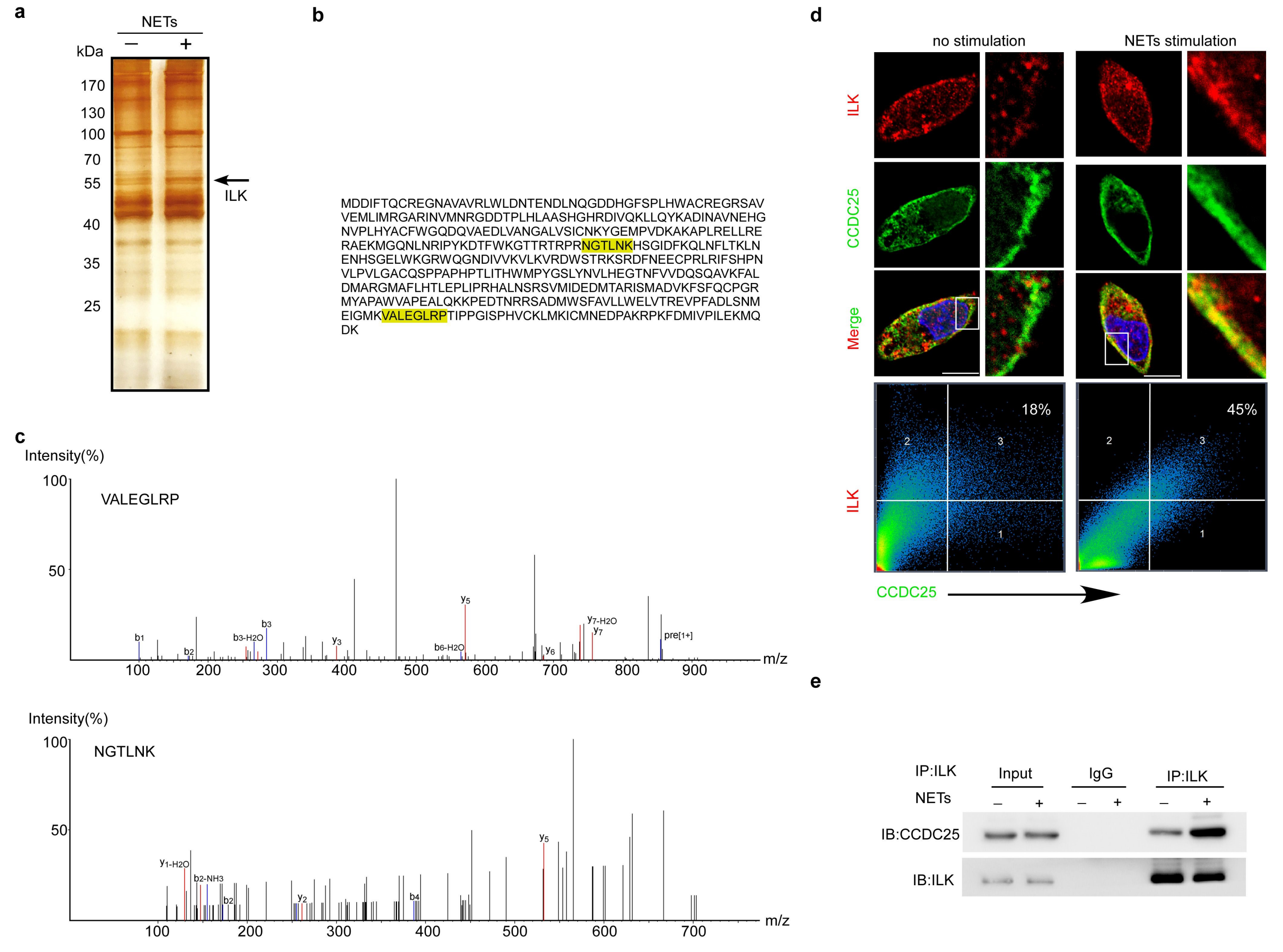


Extended Data Fig. 7 | See next page for caption.

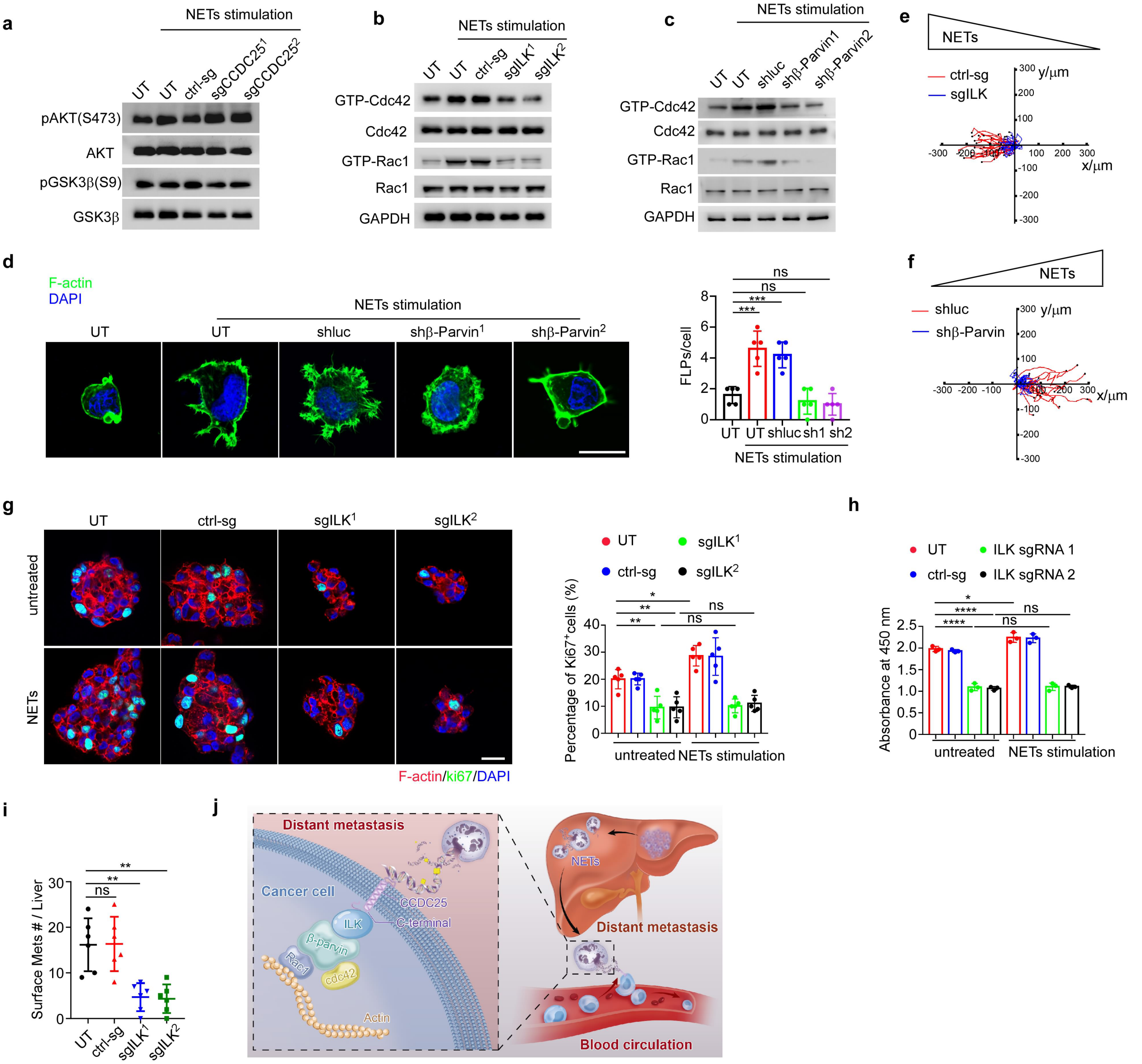
Extended Data Fig. 7 | CCDC25 is associated with poor prognosis in multiple malignant tumours. **a**, CCDC25 expression in multiple cancer types in the Human Protein Atlas database. **b**, **c**, Representative immunohistochemical staining images (left) of CCDC25 expression in primary breast cancer (*n* = 202) and liver metastasis (*n* = 17) (**b**) and in primary colon cancer (*n* = 134) and liver metastasis (*n* = 16) (**c**). Scale bars, 50 µm. Lines within the violin plots (right) mark the 25th, 50th and 75th percentiles. **P* = 0.0300, ***P* = 0.0056 as calculated by a two-sided Mann–Whitney *U*-test. **d**, Representative immunofluorescence co-staining images of CCDC25 with CK for tumour cells, CD31 for endothelial cells or α-SMA for stromal cells in human primary breast cancer. *n* = 5. **e**, Representative immunofluorescence staining images of CCDC25 in human primary breast cancer. The areas marked by the white boxes are shown magnified in the insets in the top right. *n* = 6. **f**, Representative immunofluorescence staining images for CCDC25 and H3cit in the liver metastases of patients with breast cancer. Insets as in **e**. *n* = 5. **g**, Representative immunohistochemical images for low and high CCDC25 expression in human primary breast cancer. Scale bars, 200 µm. Blue and red arrows indicate cancer cells and non-malignant cells, respectively. *n* = 573 in the low-CCDC25 group and *n* = 268 in the high-CCDC25 group. **h**, Kaplan–Meier survival curves for patients with colon cancer with high (*n* = 39) and low (*n* = 95) CCDC25 expression in the primary tumours. Comparisons are performed using a two-sided log rank test. **i**, Kaplan–Meier curves showing the overall survival of patients with breast cancer with high and low CCDC25 expression in The Cancer Genome Atlas (TCGA) breast cancer online database (*n* = 1,100); overall survival curves (*n* = 1,107) and recurrence-free survival curves (*n* = 909) of patients with lung cancer with high and low CCDC25 expression in the TCGA lung cancer online database; and overall survival curves of patients with myeloma with high and low CCDC25 expression in the Mulligan Myeloma online database (*n* = 264). The optimal survival cut points were determined by X-Tile statistical software. Comparisons were performed using a two-sided log-rank test.



Extended Data Fig. 8 | CCDC25 is a transmembrane protein and its N terminus interacts with NET-DNA. a, Transmembrane helix prediction for CCDC25. One transmembrane helix was predicted by ProtScale (https://web.expasy.org/protscale/). One confidently predicted helix (score above 0) spans residues from around 60 to 80. The N terminus (about 40 residues) and the C terminus (about 40 residues) of CCDC25 are predicted to reside at the external or cytosolic sides of the cytoplasmic membrane owing to their hydrophilicity. **b**, Membrane pellets of MDA-MB-231 lysates were resuspended in the lysis buffer or buffer containing a high salt concentration (3 M NaCl), 5 M urea, 0.2 M Na_2CO_3 (alkaline) or 0.1% Triton-X100 after centrifugation at 20,000g. The resulting lysates with the indicated treatments were separated into membrane pellets (P) and supernatants (S) by centrifugation. Immunoblotting was performed using anti-CCDC25 antibody, anti-CXCR4 antibody (positive control) or anti-ATP5b antibody (negative control). **c**, Schematics of the different CCDC25 variants. **d**, EMSA showing the binding of wild-type or mutant CCDC25 with NET-DNA. **e**–**g**, MDA-MB-231 cells were transduced with control sgRNA (Ctrl-sg) or sgRNA for *CCDC25* alone (sgCCDC25) or along with ectopic expression of full-length wild-type CCDC25 (sgCCDC25+WT) or the CCDC25 AA_{21-25} mutant (sgCCDC25+$Mutant_{(21-25)}$). **e**, The expression of indicated proteins was determined by western blot. **f**, Filopodium-like protrusions of the cells with the indicated treatments were stained with phalloidin (F-actin, red) and DAPI (nuclei, blue). Scale bars, 10 μm. *n* = 5 biologically independent experiments. Data are mean ± s.d., two-sided one-way ANOVA with Tukey test. *****P* < 0.0001, ****P* = 0.0007 and ns >0.9999. **g**, Migration tracks for the tumour cells with the indicated treatments in a chemotaxis chamber containing culture media with 5 μg ml^{-1} NET-DNA. Ctrl-sg, grey; sgCCDC25, blue; sgCCDC25+WT, green; sgCCDC25+$Mutant_{(21-25)}$, orange. **h**, CCDC25-knockout MDA-MB-231 cells with or without the ectopic expression of full-length wild-type CCDC25 or the CCDC25 AA_{21-25} mutant (M1) were injected into the spleen of NOD/SCID mice. The quantification of liver metastases is shown. *n* = 5 mice per group. Data are mean ± s.d., one-way ANOVA with Tukey test. ***P* = 0.0023 (sgCCDC25 versus ctrl-sg) and 0.0013 (sgCCDC25 + M1 versus ctrl-sg), ns = 0.9758. Data in **b**, **d**, **e**, **g** are representative of three biologically independent experiments.



Extended Data Fig. 9 | CCDC25 interacts with ILK. a, Cytosolic extracts from the HeLa cells transfected with EGFP-His-tagged-CCDC25 with (+) or without (–) NETs treatment were immunoprecipitated using anti-His antibody. Bound proteins were eluted and visualized by silver staining. A precipitated protein band of 55 kDa was submitted for mass spectrometry. **b**, The full amino-acid sequence of human ILK. The sequences in yellow are the tryptic peptides identified by liquid chromatography–mass spectrometry. **c**, Mass spectrometry analysis of the two peptides highlighted in **b**. **d**, Confocal microscopy showing the colocalization of CCDC25 with ILK in the HeLa cells transfected with His-tagged full-length CCDC25. Quantification was performed using Leica Confocal Software (fourth row). Scale bars, 10 μm. **e**, Immunoblotting of CCDC25 (top) or ILK (bottom) in the lysates (input) or immunoprecipitates (IgG or anti-ILK) of MDA-MB-231 cells stimulated with (+) or without (–) NETs treatment. Data in **a**, **d**, **e** were representative of three biologically independent experiments.



Extended Data Fig. 10 | CCDC25 interacts with ILK at its C terminus and signals through the ILK–β-parvin cascade. a, Phosphorylation of ILK substrates (AKT and GSK3β) was analysed. Whole-cell lysates of MDA-MB-231 cells transduced with two CCDC25-targeting or control sgRNAs with or without NETs stimulation at 5 μg ml⁻¹ were subjected to immunoblotting with the indicated antibodies. **b**, **c**, GTP bound or total RAC1 and CDC42 were examined in the lysates of MDA-MB-231 cells transduced with *ILK* sgRNAs (**b**) or with *PARVB* shRNAs (**c**) and stimulated with or without NETs. **d**, MDA-MB-231 cells transduced with a control shRNA or with one of two *PARVB*-shRNAs were treated with or without 5 μg ml⁻¹ NETs. The representative images of filopodium-like protrusions of the cells (stained with phalloidin (F-actin, green)) are shown on the left, and the quantification is shown on the right. Scale bars, 20 μm. $n = 5$ biologically independent experiments. Data are mean ± s.d., two-sided one-way ANOVA with Tukey test. $^{***}P = 0.0001$(UT versus UT + NETs) and 0.0007 (UT versus shluc + NETs), ns = 0.9402 (UT versus sh1 + NETs) and 0.7870 (UT versus sh2 + NETs). **e**, Migration tracks of the MDA-MB-231 cells transduced with a control sgRNA or *ILK* sgRNA in a chemotaxis chamber containing culture media with 5 μg ml⁻¹ NET-DNA. Red and blue lines denote the tracks of control and ILK-knockout tumour cells, respectively. **f**, Migration tracks of the MDA-MB-231 cells transduced with a control shRNA or with *PARVB*-shRNA in a chemotaxis chamber containing culture media with 5 μg ml⁻¹ NET-DNA. Red and blue lines denote the tracks of control and β-parvin-knockdown tumour cells. **g**, **h**, MDA-MB-231 cells transduced with a control sgRNA or two sgRNAs for *ILK* were treated with or without 5 μg ml⁻¹ NET-DNA. **g**, Left, representative images of immunofluorescence staining for ki67 (green) and F-actin (red) in MDA-MB-231 cells in a 3D culture system. Scale bars, 20 μm. Right, quantification of the ki67-positive tumour cells. $n = 5$ biologically independent experiments. Data are mean ± s.d., two-sided one-way ANOVA with Tukey test. $^{**}P = 0.0050$ (UT versus sgILK¹) and 0.0055 (UT versus sgILK²), $^{*}P = 0.0328$, ns >0.9999 (sgILK¹ versus sgILK¹ + NETs) and 0.9990 (sgILK² versus sgILK² + NETs). **h**, Cell proliferation was assessed by the CCK-8 assay in a 2D culture system. $n = 3$ biologically independent experiments, Data are mean ± s.d., two-sided one-way ANOVA with Tukey test. $^{*}P = 0.0105$, $^{****}P < 0.0001$, ns >0.9999 (ILK sgRNA1 versus ILK sgRNA1 + NETs) and = 0.9969 (ILK sgRNA2 versus ILK sgRNA2 + NETs). **i**, MDA-MB-231 cells that were untreated or transduced with a control or with one of two *ILK* sgRNAs were intrasplenically injected into NOD/SCID mice, and liver metastatic nodules were counted 30 days after injection. $n = 6$ mice per group. Data are mean ± s.d., ns >0.9999 and $^{**}P = 0.0021$ (UT versus sgILK¹) and 0.0016 (UT versus sgILK²), determined by a two-sided one-way ANOVA with Tukey test. **j**, Schematics highlighting the major findings of this study. Data in **a**–**c**, **e**, **f** are representative of three independent experiments.

nature research

Corresponding author(s): Erwei Song, Shicheng Su

Last updated by author(s): Mar 12, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a | Confirmed

☐ ☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement

☐ ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly

☐ ☒ The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.

☒ ☐ A description of all covariates tested

☐ ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons

☐ ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)

☐ ☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted
Give P values as exact values whenever suitable.

☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings

☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes

☐ ☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Fluorescent images: Laser scanning confocal microscopy(LSM780 or 800, Zeiss). Flow cytometry: Accuri C6, BD.
Data analysis	Data representation, Kaplan-Meier prognostic analysis and statistical analysis: Graphpad Prism 7.0; Laser scanning confocal microscopy: Imaris 9.0;ZEN 2012; IVIS Lumina Imaging: Living Image software ver. 3.0.; Flow cytometry data analysis: FlowJo (versions 7.6); GeneSpring GX v12.1 software package (Agilent Technologies), GSEA software (version 4.0.3),Image J software and Chemotaxis and Migration Tool,Bio-layer interferometry.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The authors declare that the data supporting the findings of this study are available within the paper and its supplementary information files or from the corresponding author upon reasonable request. Source data for the RNA microarray shown in Extended Data Fig. 10 have been deposited in the GEO repository under accession code GSE142281.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For clinical sample analysis, sample size was determined on the basis of similar research reported in the literature. For in vitro experiments, the sample size was determined based on pilot experiments or previous studies. Sample size of animal experiments was estimated on the basis of similar research reported in the literature. All experiments included at least 3 independent experiments. The number of independent experiment was indicated in each figure legend.
Data exclusions	No data were excluded.
Replication	For each experiments the number of biological independent animal/sample/patient is reported in the figure legend. In vitro studies are represented at least 3 independent reproducible studies. Animals studies represent at least 5 independent mice. Studies using human tumor slices represent reproducible observations from independent cohorts of breast and colon cancer patients.
Randomization	Animals were allocated randomly to each treatment group. Different treatment groups were processed identically, and animals in different treatment groups were exposed to the same environment.
Blinding	Blinding was used in clinical sample analysis including immunohistochemical/immunofluorescent staining and quantification.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study	
☐	☒	Antibodies
☐	☒	Eukaryotic cell lines
☒	☐	Palaeontology
☐	☒	Animals and other organisms
☐	☒	Human research participants
☒	☐	Clinical data

Methods

n/a	Involved in the study	
☒	☐	ChIP-seq
☐	☒	Flow cytometry
☒	☐	MRI-based neuroimaging

Antibodies

Antibodies used

All the antibodies are from commercial sources and have been validated by the vendors and their validation data are available on the manufacturer's website. Antibody used for immunofluorescence (IF), enzyme linked immunosorbent assay (ELISA), immunoprecipitation (IP), immunoblotting (IB), immunohistochemistry (IHC) and Flow cytometry (FC) with their respective catalogue number and vendor is mentioned below.

H3Cit ab5103 Abcam IF:100 IB:1:1000
MPO AF3667 R&D IF: 10 ug/ml
Cytokeratin GTX27753 Genetex IF:1:50
8OHdG GTX41980 GeneTex IF:1:50
Flag 14793 CST IF:1:100
Alexa Fluor 488 Phalloidin A12379 Invitrogen IF:165 nM
Alexa Fluor 555 Phalloidin A34055 Invitrogen IF:165 nM
CM-Dil C7000, Thermo Fisher IF: 1uM
CFSE C34554, Thermo Fisher Cell label: 0.5 uM
celltracker Orange CMTMR Dye C2927, Thermo Fisher Cell label: 5 uM
Hoechst 33342 H21492, Thermo Fisher Cell label: 5 ug/ml
MPO monoclonal antibody 0400-0002 ABD Serotec ELISA: 5 ug/ml
Flag-tag 637309 BioLegend FC:1:100
CCDC25 21209-1-AP ProteinTech IB:1:1000, IHC:1:50
CCDC25 PA5-54735 Invitrogen IF:1:50

CDC25 sc-515201 Santa Cruz IF:1:50
Ki67 ab16667 Abcam IF:1:50
CD31 ab9498 Abcam IF:1:50
CD31 ab56299 Abcam IF:1:50
PDGFRβ ab69506 Abcam IF:1:50
α-SMA MAB1420 R&D IF:1:50
ILK 3862 CST IB:1:1000 IF: 1:100
ILK 3856 CST IP:1:50
His-tag MA1-21315 Thermo Fisher IB:1:1000, IF:1:100, IP: 1:100
β-Parvin 14463-1-AP ProteinTech IB:1:1000
α-Parvin 4026 CST IB:1:1000
Pinch 11890 CST IB:1:1000
CXCR4 60042-1-Ig ProteinTech IB:1:1000
ATP5b 17247-1-AP ProteinTech IB:1:1000
Clathrin ab2731 Abcam IB:1:1000
H2AX ab11175 Abcam IB:1:1000
COXIV 11242-1-AP ProteinTech IB:1:1000
GM130 11308-1-AP ProteinTech IB:1:1000
Calreticulin 11242-1-AP ProteinTech IB:1:500
GAPDH HRP-60004 ProteinTech IB:1:10000

Validation	All antibodies used in this study were obtained from commercial sources and validated according to manufacturers' instruction.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Murine E0771 breast cancer cell line was purchased from CH3 Biosystems (New York, USA). MDA-MB-231, MCF-7, HCT116 HEK293T, HeLa and 4T1 cells were obtained from American Type Culture Collection (ATCC).
Authentication	All the cell lines were authenticated by short tandem repeat profiling prior to use.
Mycoplasma contamination	All the cell lines were tested negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used in this study.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	PAD4flox/flox mice, Ddx4-Cre and B6.FVB-Tg (MMTV-PyMT) mice were purchased from the Jackson Laboratory. CCDC25 knockout mice were constructed by Shanghai Model Organisms Center. 6-8 week-old female C57BL/6 mice, BALB/c mice and NOD/SCID mice were maintained at the Animal Experiment Center of Sun-Yat-Sen University
Wild animals	No wild animals were used.
Field-collected samples	For the intrasplenical injection of cancer cells, the mouse liver tissues were isolated after 30 days; for the tail-vein injection model, the lung tissues were isolated after 30 days. For the orthotopic transplantations, the primary tumor, liver and lung tissues were isolated in the indicated time shown in figure legend. All the samples were isolated and then fixed in polyformaldehyde for paraffin embedding later.
Ethics oversight	All mouse experiments were reviewed and approved by the Animal Care and Use Committee of Sun-Yat-Sen University.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Immunofluorescent staining for NETs was performed in the tissues of primary tumors (461 cases) , liver metastasis (20 cases), lung metastasis (23 cases), bone metastasis (33 cases) and brain metastasis (7 cases) of breast cancer patients who were collected from Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University (Guangzhou, China) between 2007 and 2016. The serum and plasma samples were collected from breast cancer patients admitted to Sun Yat-Sen Memorial Hospital between 2011 and 2019. Moreover, immunofluorescent staining for NETs was also performed in the tissues of primary tumors (130 cases), liver metastasis (16 cases), lung metastasis (12 cases), bone metastasis (3 cases) and brain metastasis (5 cases) of colon cancer patients admitted to Sun Yat-Sen Memorial Hospital. Immunohistochemical staining for CCDC25 was performed in the primary tumor tissues (202 cases) and liver metastasis (17 cases) of breast cancer patients and in the primary tumor tissues (134 cases) and liver metastasis (16 cases) of colon cancer patients. Additionally, immunohistochemical staining for CCDC25 was performed in 841 breast cancer samples and 134 colon cancer samples from Sun Yat-Sen Memorial Hospital.

Recruitment	All samples were collected from the patients who had provided informed consent at Sun-Yat-Sen Memorial Hospital, and all the related procedures were performed with the approval of the internal review and ethics board of Sun-Yat-Sen Memorial Hospital.
Ethics oversight	All the related procedures were performed with the approval of the internal review and ethics board of Sun-Yat-Sen Memorial Hospital.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Hela cells were dissociated by 0.25% trypsin-EDTA, harvested by centrifugation.
Instrument	Accuri C6
Software	FlowJo (versions 7.6).
Cell population abundance	Purity of FACS-sorted samples was analysed by flow cytometry. Purity of the samples was >95%.
Gating strategy	Starting cells were gated by FSC/SSC gates. Gates indicating boundaries between "positive" and "negative" are according to the isotype staining. Expression of indicated proteins were checked on these populations as indicated in the figures and figure legends

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Base-pair conformational switch modulates miR-34a targeting of Sirt1 mRNA

https://doi.org/10.1038/s41586-020-2336-3

Received: 28 September 2018

Accepted: 12 March 2020

Lorenzo Baronti[1], Ileana Guzzetti[1,6], Parisa Ebrahimi[2,6], Sarah Friebe Sandoz[1,6], Emilie Steiner[1,4,6], Judith Schlagnitweit[1], Bastian Fromm[3], Luis Silva[1], Carolina Fontana[1,5], Alan A. Chen[2] & Katja Petzold[1]✉

MicroRNAs (miRNAs) regulate the levels of translation of messenger RNAs (mRNAs). At present, the major parameter that can explain the selection of the target mRNA and the efficiency of translation repression is the base pairing between the 'seed' region of the miRNA and its counterpart mRNA[1]. Here we use $R_{1\rho}$ relaxation-dispersion nuclear magnetic resonance[2] and molecular simulations[3] to reveal a dynamic switch—based on the rearrangement of a single base pair in the miRNA–mRNA duplex—that elongates a weak five-base-pair seed to a complete seven-base-pair seed. This switch also causes coaxial stacking of the seed and supplementary helix fitting into human Argonaute 2 protein (Ago2), reminiscent of an active state in prokaryotic Ago[4,5]. Stabilizing this transient state leads to enhanced repression of the target mRNA in cells, revealing the importance of this miRNA–mRNA structure. Our observations tie together previous findings regarding the stepwise miRNA targeting process from an initial 'screening' state to an 'active' state, and unveil the role of the RNA duplex beyond the seed in Ago2.

MicroRNAs—non-coding RNA molecules—regulate gene expression by targeting mRNAs. Each mature miRNA of roughly 22 nucleotides is bound to one Argonaute protein (Ago1 to Ago4 in humans), forming an RNA-induced silencing complex (RISC). In the RISC, nucleotides 2–6 of the guide miRNA (g2–g6) are prearranged to recognize mRNA targets through Watson–Crick base pairing[6–8] in the seed (Fig. 1a, b). This base-pair complementarity (involving up to g2–g8) largely determines RISC activity[9,10]: for example, complementarity involving just g2–g6 (a 5-mer) is rejected as unspecific. In human Ago2 (hereafter, Ago2 refers to human Ago2 unless specified otherwise), sites with prolonged base pairing, using at least g2–g7 base-pairing (a 6-mer or larger), can override the checkpoint imposed by Ago2's flexible helix-7 (ref. [11]) and induce a conformational transition in Ago2, allowing extended 3′-pairing of the RNA[12]. However, bioinformatics analysis of validated miRNA–mRNA pairs cannot discern sequence determinants in this region, beyond a preference for forming bulges[13]. Moreover, X-ray structures of ternary complexes are unable to resolve the central region of the duplex, supporting the idea of its flexibility[14]. In vitro biochemical studies[15] showed that mismatches in this region contribute little to target binding affinity but can impair catalytic cleavage of short interfering RNAs (siRNAs) in *Drosophila* Ago2. This implied that the dynamics of the central RNA bases are essential for the fate of target mRNAs; however, the precise nature of the guide–target interaction beyond the seed region remained unclear.

Here we use nuclear magnetic resonance (NMR) to observe the dynamic process underlying miRNA–mRNA targeting. To elucidate the effects of the conformational transition on RISC function, we combine these measurements with molecular simulations and dual-luciferase reporter (DLR) assays in human cells.

We study hsa-miR-34a-5p ('miR-34a'), part of the evolutionarily conserved miR-34/449 family of miRNAs[16], which targets the mRNA encoding silent information regulator 1 (Sirt1)—a p53-deacetylating enzyme—in a tumour-suppressive feedback loop. Using $R_{1\rho}$ NMR relaxation dispersion, we show that the weak so-called 7-mer–A1 seed of the miR-34a–mSirt1 duplex (Fig. 1a, b) is in equilibrium with a transient and low-populated excited state that results in an 8-mer seed with a G:U base-pair at its 3′-end. The extended seed alters the topology of the duplex by shifting the bending angle between the seed and the 3′-helix in the RISC, as shown by simulations. In a cell-based assay, a structural mimic of the extended seed produces a roughly two-fold increase in target downregulation. Our data suggest a model whereby RISC undergoes a structural transition mediated by RNA dynamics: the RISC first screens targets for correct seed pairing, than transitions into an active complex, releasing the miR-34a 3′-end which is allowed to fully bind the Sirt1 mRNA in the compensatory region.

Seed dynamics of miR-34a–mSirt1 binding site

Given that an RISC recognizes thousands of distinct binding sites in its target mRNA, with no apparent sequence preference beside the seed, we hypothesize that miRNA–mRNA pairs possess distinct conformational characteristics in the central bulge, facilitating their accommodation within Ago2.

---

[1]Department of Medical Biochemistry and Biophysics, Karolinska Institute, Stockholm, Sweden. [2]Department of Chemistry and RNA Institute, University at Albany, State University of New York, Albany, NY, USA. [3]Science for Life Laboratory, Department of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, Stockholm, Sweden. [4]Present address: Laboratoire SABNP, UMR INSERM 1204, Université d'Evry-Val-d'Essonne–Université Paris-Saclay, Evry, France. [5]Present address: Departamento de Química del Litoral, Facultad de Química and CENUR Litoral Norte, Universidad de la República, Paysandú, Uruguay. [6]These authors contributed equally: Ileana Guzzetti, Parisa Ebrahimi, Sarah Friebe Sandoz, Emilie Steiner. ✉e-mail: katja.petzold@ki.se

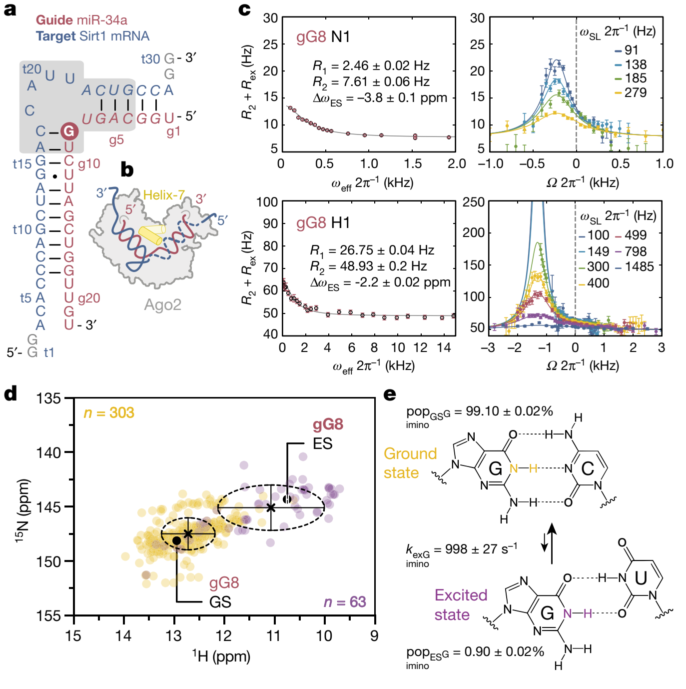


Fig. 1 | Conformational dynamics in the seed of miR-34a–Sirt1 mRNA. **a**, Secondary structure of the miR-34a–mSirt1 duplex determined by NMR. The seed (g2–g6 bound to t23–t27) comprises five base-pairs. The grey box indicates nucleotides selected to generate the reduced construct for $R_{1\rho}$ relaxation-dispersion measurements. **b**, Sketch of human Ago2 accommodating the miR-34a–mSirt1 duplex. Helix-7 is part of Ago2 and is shown with yellow cylinders. **c**, ^{15}N and ^{1}H $R_{1\rho}$ individual relaxation-dispersion profiles of gG8N1 and gG8H1, revealing the single-base-pair switching of gG8:C to gG8:U (circled in **a**). R_1 and R_2 are the longitudinal and transverse relaxation rates, respectively; R_{ex} is the exchange contribution to the relaxation rate; $\omega_{eff}\,2\pi^{-1}$ is the effective measured spinlock power; $\Omega\,2\pi^{-1}$ is the offset; $\Delta\omega_{ES}$ is the chemical shift, a structural parameter of the ground-state-to-excited-state transition; $\omega_{SL}\,2\pi^{-1}$ is the measured spinlock power. Error bars represent one standard deviation (s.d.), derived from Monte Carlo simulations of experimental uncertainty (see Supplementary Methods). **d**, Chemical-shift distribution for ^{1}H1–^{15}N1 moieties of guanosines in G:C (yellow) or G:U (purple) base pairs from the BMRB[20]. Crosses indicate average chemical shifts for G:C and G:U; dashed ellipses show 1 s.d.; black dots indicate chemical shifts for gG8 in the ground state (GS) and relaxation-dispersion-derived excited state (ES). **e**, The G:C to G:U base-pair switch, highlighting the guanosine ^{1}H1–^{15}N1 (imino-global fit, one-sided *F*-test, $n = 1$) groups in the ground state (yellow) and excited state (purple). Errors represent 1 s.d. derived from Monte Carlo simulations of experimental uncertainty (see Supplementary Methods).

First, we solved the secondary structure of miR-34a bound to the validated target site in Sirt1 mRNA (miR-34a–mSirt1 duplex)[17] by NMR (Fig. 1a, Supplementary Fig. 2 and Supplementary Discussion section 2). The overall fold confirmed the secondary structure predicted using MC-Fold (Fig. 1a and Extended Data Fig. 1): the five-nucleotide seed constitutes an A-form 5′-helix between the gG2:tC27 and gG6:tC23 base pairs; meanwhile gG8:tC17 and gG18:tC7 form a 3′-helix containing a wobble gU11:tG14 base pair ('t' refers to the target mRNA). These two helices are separated by a four-nucleotide asymmetric bulge on the mSirt1 side, comprising tC18–tU21 (Figs. 1a, 2a).

To study the structure and dynamics of the bulge, we designed a shortened hairpin construct (miR-34a–mSirt1 bulge) containing the four-nucleotide bulge and enclosing regions (Fig. 1a, grey box, and Fig. 2a). The correct fold was confirmed by a chemical-shift comparison of the shared residues (Supplementary Fig. 2e–h). The intrinsic flexibility of the miR-34a–mSirt1 complex precluded a traditional NMR tertiary-structure calculation with a single, static conformation. Therefore, we used an NMR-informed computational approach and computed the RNA's conformational ensemble using replica-exchange molecular dynamics (REMD) simulations, constraining the base-pairing determined from imino ^{1}H–^{1}H nuclear Overhauser effects (NOEs) data derived by NMR (Supplementary Fig. 3). We varied the temperature in the simulations in order to explore the RNA conformations that fulfil the experimental constraints, resulting in an ensemble of 153 structures. One representative structure from the ensemble is shown in Fig. 2d, with the relative stem-to-stem angle distribution shown in Fig. 2g (left).

Although classified as a 7-mer–A1 binding site by prediction tools (for example, Targetscan[18]), we found that the miR-34a–mSirt1 duplex and the reduced construct represent a less stable structure: NMR shows that the stability of the gU7:tA22 closing base pair at the 3′-end of the seed is substantially reduced (Fig. 2a, Supplementary Information Figs. S2a, S3a and Supplementary Discussion sections 2, 3). We suggest, therefore, that weak pairing at position 7 might explain previously observed sequence-specific differences in the binding affinity between RISC and target[12]. In agreement with nearest-neighbour models for A:U closing hairpins[19], we propose that 6-mer/7-mer–A1 seeds ending with closing A:U base pairs at position 7 might not suffice for stable displacement of helix-7 of Ago2, resulting in much lower binding affinities, closer to the predicted affinity of the 6-mer.

To assess the base-pair dynamics, we carried out ^{15}N, ^{13}C and ^{1}H $R_{1\rho}$ NMR relaxation-dispersion experiments[2]. ^{1}H–^{15}N NMR of gG8H1 and gG8N1, and ^{13}C NMR of gG8C8, tU21C6, tC17C1′, tU20C1′, tA19C8, tA19C2 and tA22C8, revealed a global exchange process. In this process, the base pair gG8:tC17 interconverts from the most stable structure, the ground state, to a low-populated excited state. The exchange-rate constant (based on ^{1}H1–^{15}N1) for gG8 ($k_{exGimino}$) is 998 ± 27 s^{-1}, with an excited-state population ($pop_{ESGimino}$) of 0.90 ± 0.02; the global k_{ex} (k_{exG}) is 1,008 ± 12 s^{-1}, with an excited-state population (pop_{ESG}) of 0.90 ± 0.01% (Fig. 1c, e, Supplementary Fig. S6c and Supplementary Data S1 Tab 1). Most importantly, we obtained the individual chemical-shift difference between the ground and excited states, $\Delta\omega_{ES} = \Omega_{ES} - \Omega_{GS}$—describing the structure of the excited state—by measuring ^{1}H ($\Delta\omega_{ES} = -2.20 \pm 0.02$ ppm) and ^{15}N ($\Delta\omega_{ES} = -3.8 \pm 0.1$ ppm) in $R_{1\rho}$ relaxation-dispersion data sets. This approach allows us to infer that chemical shifts in the gG8 excited state reside in a region of the ^{1}H–^{15}N heteronuclear single quantum coherence (HSQC) spectrum that is a signature for G:U wobble-base-paired guanosines. This was validated by querying the Biological Magnetic Resonance Bank (BMRB)[20] for ^{1}H1–^{15}N1 chemical shifts of G:U base-paired Gs in RNA-only entries and comparing them with the G:C distribution (Fig. 1d, e, Supplementary Fig. 6c and Supplementary Data S1 Tab 1).

Base-pair switch alters the complex topology

When analysing the MC-Fold[21] output for alternative secondary structures that could fulfil the NMR-derived model (Fig. 1e), we found that a switch in base-pairing partner from gG8:tC17 to gG8:tU21 occurs within the third most energetically favourable structure (Extended Data Fig. 1a, b). To characterize the nature of this process, we carried out additional ^{13}C $R_{1\rho}$ relaxation-dispersion experiments on aromatic C2/6/8 and sugar C1′ nuclei—known reporters of sugar pucker, stacking and base pairing. The additional, individually fitted nuclei resulted in an exchange process with average parameters of k_{EX} = 1,371 Hz and pop_{ES} = 1.9%, similar to the global fit obtained with ^{1}H, ^{13}C and ^{15}N data sets (Supplementary Figs. 6, 7 and Supplementary Data S1 Tabs 1–3). On the basis of known correlations between our measured ^{13}C $R_{1\rho}$ $\Delta\omega$ values and structural propensities[22–24], we propose a refined secondary structure of the excited state (Fig. 2a).

To derive a three-dimensional structural model of the excited state, we carried out high-temperature REMD simulations of the ground state, restraining five experimentally determined base pairs (Supplementary Data S1 Tab 11)[3]. We identified a putative conformation of the

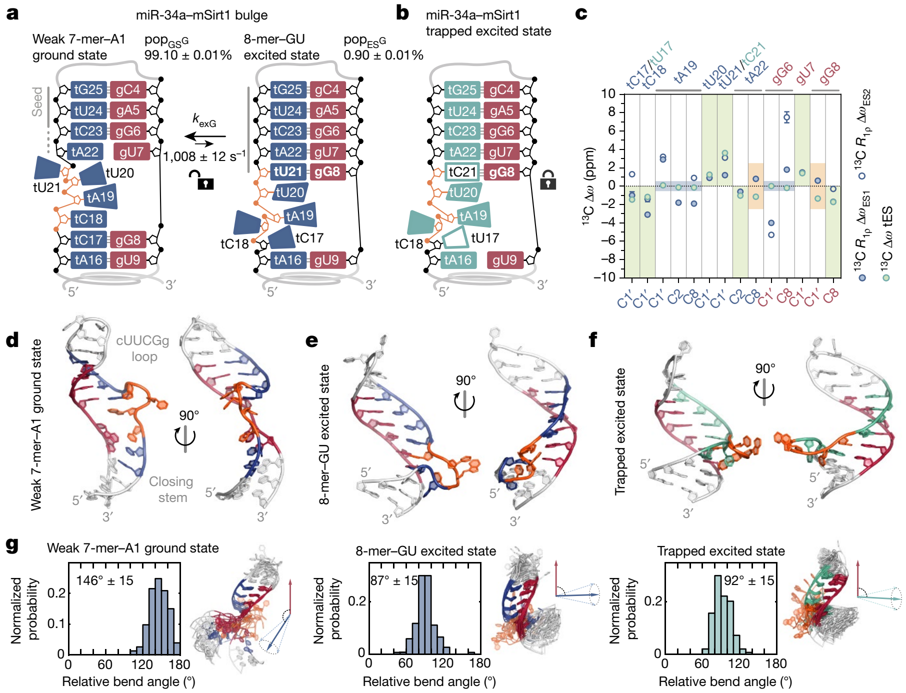


Fig. 2 | Structure and conformation of the excited state of miR-34a–mSirt1. **a**, Secondary structures of the bulge region (the grey area in Fig. 1a). Left, the ground state as solved by NMR; right, the excited state, resulting from $R_{1\rho}$ relaxation-dispersion-derived chemical shifts (gG8:tC17 to gG8:tU21). **b**, Stabilization of the excited-state conformation by isosteric two-point substitution of tC17 with tU17 and tU21 with C21 (trapped excited state). This secondary structure was solved by NMR. **c**, MCSF analysis of the trapped excited state (tES) validates the excited-state model (green shading). Expected perturbations are observed at the sites of modification (orange shading); tA19 and gG6 (blue shading) are explained in Supplementary Discussion 5 and Extended Data Fig. 5. Individually fitted ^{13}C $R_{1\rho}$ relaxation-dispersion-derived Δω values are in blue, with filled dots for excited state 1 and hollow dots for excited state 2, for three-state-fitting data sets. Error bars for $R_{1\rho}$ relaxation-dispersion-derived Δω values represent 1 s.d. from fitting (see Supplementary Methods). **d**–**f**, Representative conformations from NMR-informed REMD of the 7-mer–A1 ground state (**d**), the 8-mer–GU excited state (**e**) and the trapped excited state (**f**). **g**, Interhelical bend-angle distributions for the ground state (cluster size $n = 153$), excited state (cluster size $n = 210$) and trapped excited state (cluster size $n = 222$). Means ± s.d. of angle distribution are derived from REMD.

excited state as a cluster within simulations of the ground state, with gG8:tU21 being base paired to gG8:tC17. We sampled the excited-state conformer, restraining gU9:tA16 and gG6:tC23. Addition of magnesium ions, experimentally and in simulations, had no effect (Extended Data Fig. 2 and Supplementary Fig. 6). As for the ground state, we show one representative structure from the excited-state ensemble (210 structures) in Fig. 2e. The topology of the excited state is altered compared with the ground state, indicated by a stem-to-stem coaxial stack that results in an angle distribution peaking around 90° (Fig. 2g, middle).

To experimentally validate the candidate excited-state structure, we used the NMR mutate-and-chemical-shift-fingerprint (MCSF) approach[22], where a substitution or chemical modification is used to trap the proposed excited state. Chemical shifts are then compared between the trapped excited-state and the $R_{1\rho}$ relaxation-dispersion-derived data. We introduced a two-point isosteric substitution in the bulge construct, swapping tC17 with tU17 and tU21 with tC21. This promotes the repositioning of gG8 to the seed 5′-helix, base paired with tC21 (we name this the 'miR-34a–mSirt1 trapped excited state'), without affecting the overall binding affinity. We determined the secondary structure of the trapped excited state by NMR (Fig. 2b and Supplementary Fig. 4) and used imino ^{1}H-^{1}H NOEs as sparse constraints to calculate structural ensembles via REMD (Fig. 2f). As expected, the trapped excited state forms an additional gG8:tC21 base pair that elongates the seed 5′-helix, resulting in identical base-pairing patterns and interhelical bending angles to those in the excited state (Fig. 2g, right).

The MCSF showed remarkable agreement for C1′s, tA22C2 and gG8C8 (Fig. 2c, green) confirming that the trapped excited state well represents the overall topology of the excited state modelled from $R_{1\rho}$ relaxation-dispersion data. The sugar puckers measured by $^{3}J_{H1'-H2'}$ for tU20, tU21 (dominant C2′-endo) and tC18 (dominant C3′-endo)—which were expected on the basis of $R_{1\rho}$ relaxation dispersion to interconvert to their opposite configuration in the excited state (Fig. 2a)—were successfully recapitulated in the trapped excited state (Fig. 2b). Furthermore, coaxial stacking between the two helices is validated by tA22H8/C8, tA16H8/C8 and gG8H8/C8 chemically shifting to a region that is characteristic of nucleotides embedded in the uninterrupted A-form helix[23] (Extended Data Fig. 3 and Supplementary Fig. 5a).

Inconsistencies observed for tA22C8 and gG8C1′ are a consequence of the substitution (Fig. 2c, orange, and Extended Data Fig. 3). Data sets for tA19 and gG6 reveal the presence of a second, thermodynamically similar excited state (ES2 in Fig. 2c, blue). However, this conformation could not be trapped experimentally and is discussed in Supplementary Discussion 5 and Extended Data Fig. 3. Interestingly, when probing the dynamics of the trapped excited state in relaxation-dispersion experiments, we detected no exchange with alternative conformations during the timescale probed (Supplementary Figs. 8, 9 and Supplementary Data S1 Tab 3).

In summary, our results show that the miR-34a–mSirt1 binding site is in equilibrium between a high-populated but weak 7-mer–A1 ground state and a low-populated 8-mer–GU seed-elongating excited state, where position 8 is occupied by a G:U base pair, a motif seen previously for miR-48[25,26]. During the ground-state to excited-state switch, both $R_{1\rho}$ relaxation-dispersion data and REMD indicate rearrangement of the bulge and stacking of the two helices.

Functional relevance of 8-mer–GU excited state

We compared wild-type miR-34a–mSirt1 and the miR-34a–mSirt1 trapped excited state by measuring thermal stability followed by

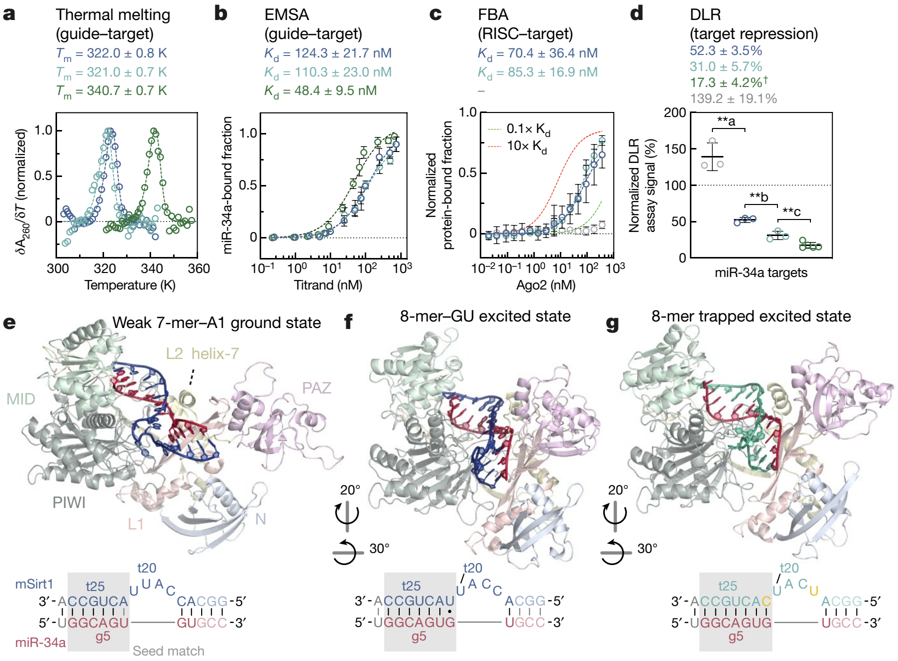


Fig. 3 | Biophysical and functional characterization of wild-type and trapped excited state miR-34a–mSirt1 duplexes. a–c, The wild-type miR-34a–mSirt1 (blue) and the complex in its trapped excited state (turquoise) show comparable stability as indicated by their equivalent melting temperature (T_m) (**a**), binding affinity (K_d) (**b**) and filter-binding assay (FBA; **c**) values. The green curves show the perfect complement for purposes of comparison of the RNAi pathway downregulation efficiency and stability. The melting temperatures (**b**) were obtained by thermal denaturation monitored by ultraviolet absorption at 260 nm (A_{260}); shown are plots from single technical replicates. T_m values are presented as means ± of fitted T_m values from three individual technical replicates. K_d values (**b**) were obtained by EMSA; means are at the plot centres, and error bars represent 1 s.d. from three independent replicates. Fitted K_d values are presented with a confidence interval of 95% as an estimation of the experimental error. For FBA analysis of binding to the miR-34a-loaded RISC (**c**), means are at the plot centres for each data point, and error bars represent 1 s.d. from three independent replicates. See Extended Data Fig. 4, Extended Data Table 1a–c and Supplementary Methods for further details. **d**, DLR assays reveal a roughly two-fold increase in miR-34a-mediated downregulation for the trapped excited state (turquoise, $n = 3$) with respect to the wild-type (blue, $n = 3$). Grey, scrambled negative control, $n = 3$. Green, the highest level of downregulation (siRNA-type), $n = 5$ (performed independently). *P*-values: **a = 0.0015, **b = 0.0054, **c = 0.0076 (** indicates $P < 0.01$, unpaired, two-tailed *t*-test), The centre line shows the mean and error bars represent 1 s.d. from independent replicates (see Supplementary Table 12). **e–g**, Slow-growth simulated RNA structures bound to Ago2 (Protein Data Bank (PDB, https://www.rcsb.org) code 4W5O)[12]. **e**, The ground-state conformation (Fig. 2a, d), orients the compensatory region towards the PAZ domain. **f**, The excited-state conformation (Fig. 2a, e), with coaxial stacking of the helices, orients it towards the N-PIWI domain. **g**, The trapped excited state recapitulates the excited state in **f**.

ultraviolet absorption; RNA–RNA binding affinity by electrophoretic mobility shift assay (EMSA); and RISC–target affinity by filter-binding assay (FBA) of miR-34a-loaded Ago2 (Fig. 3a–c). We found that the melting temperature (T_m) and dissociation constant (K_d) were unchanged (Fig. 3a, b and Extended Data Fig. 4), showing that the substitution does not affect the duplex stability in vitro. Similarly, the binding affinity of miR-34a-loaded Ago2 for the target RNAs in FBA is the same within error (Fig. 3c). Next, we asked whether the two binding sites, despite their similar stabilities, produce different degrees of target downregulation in cells. DLR assays in HEK 293T cells of miR-34a co-transfected with the wild-type weak 7-mer–A1 results in 52.3 ± 3.5% downregulation, as previously reported[17] (Fig. 3d, blue), while the trapped 8-mer excited state leads to 31.0 ± 5.7% downregulation (Fig. 3d, turquoise), showing that the two-point substitution that traps the excited state causes a roughly two-fold increase in target downregulation. Taken together, the FBA and DLR assays suggest that, for stably bound 3′-paired targets, the binding affinity cannot fully explain the observed biological data.

This difference prompted us to compute the RNA structure in the context of RISC[12]. We used slow-growth simulations to test whether the calculated ensembles of the miR-34a–mSirt1 bulge ground state, excited state and trapped excited state (Fig. 2d–f) could be accommodated in the Ago2 binding site. Starting from the crystal structures[12], we replaced the visible crystallographic A-form seed helix with conformations from the miR-34a–mSirt1 bulge ground-state, excited-state and trapped excited-state ensembles and aligned them with the seed of the cocrystal (Extended Data Fig. 5).

The resulting simulated ternary complexes are shown in Fig. 3e–g. The ground-state ensemble samples the 3′-helix of the miRNA–mRNA complex within the PAZ domain (Fig. 3e), where the miRNA is bound before target binding[6,7,12,27]. By contrast, the 8-mer–GU excited-state conformation adopts a global bend angle that stacks the 3′-helix coaxially with the seed and favours binding along the PIWI-N domains (Fig. 3f), also recapitulated in the trapped excited state (Fig. 3g).

Although only small conformational changes in the crystal structure of Ago2 are needed to bind the miRNA–mRNA complex in the ground state conformation, accommodating the excited state conformation requires pivoting of the PAZ domain (Fig. 4e and Supplementary Video 1), consistent with prior studies, in which simulations identified these PAZ-domain movements as leading to more 'open' Ago2 conformations[28]. Intriguingly, the slow-growth induced-fit conformation of Ago2 bound to the excited state is reminiscent of the binding modes observed for DNA-bound prokaryotic Ago ternary complexes[4,5] (Extended Data Fig. 6), suggesting that Ago2 undergoes structural changes during target recognition and downregulation activity.

We therefore performed a sequence search for other instances of ground-state to excited-state transitions in the 28,653 isoforms of 19,432 human protein-coding genes (specifically, in their 3′-untranslated regions). Requiring a minimal 6-mer–A1 seed resulted in 3,269 predicted target sites for miR-34a (Fig. 4a). Using MC-Fold[21], we then

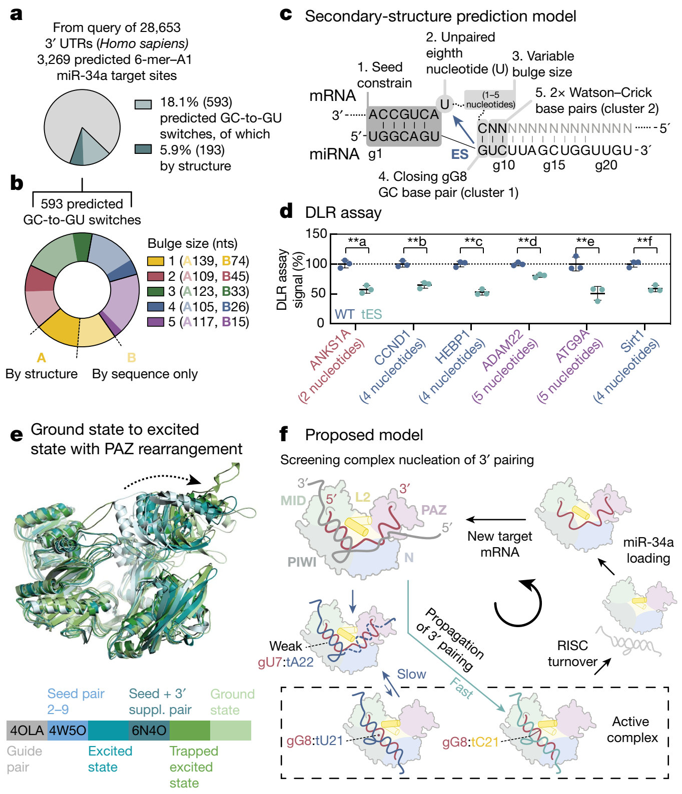


Fig. 4 | Proposed mechanism of downregulation for GC–GU switches in miR-34a-loaded RISC. a, Predicted miR-34a targets in human 3′-UTRs (grey), showing those predicted by sequence (18.1%) and by structure (5.9%) to experience the GC-to-GU switch. **b**, Distribution of bulge sizes (predicted by sequence, lighter colours, or structure, darker colours). **c**, Model of criteria used to search for GC-to-GU switches. **d**, DLR assay for repression of five target mRNAs with respective trapped excited states. All data sets were normalized internally and to the wild type for comparison. The experiment for Sirt1 was performed independently. *P*-values: **a = 0.001324, **b = 0.001112, **c = 0.000253, **d = 0.000935, **e = 0.006454, **f = 0.000646 (**$P < 0.01$, unpaired, two-tailed multiple *t*-test, $n = 3$). The centre line shows the mean; error bars represent 1 s.d from independent replicates (see Supplementary Table 12). **e**, Overlay of X-ray structures and slow-growth simulations. The code at bottom explains the colour of each structure. **f**, Proposed mechanism for the miRNA–mRNA switch in the RISC from a 'screening' into an 'active' state. Domains of Ago2 are indicated in the top left structure. See Discussion for more details.

carried out a sequence search and secondary-structure prediction for the ground-state-to-excited-state switch motif, resulting in bulge sizes from one nucleotide (139 and 74 representatives from sequence search and secondary-structure prediction, respectively), two nucleotides (109 and 45), three nucleotides (123 and 33), four nucleotides (105 and 26) and five nucleotides (117 and 15), respectively. In a more stringent cluster, with three Watson–Crick base pairs following the bulge, we identified 22 targets (Fig. 4c). We selected five different mRNA targets for further investigation in DLR assays in HEK 293T cells (HEBP1, ADAM22, ATG9A, ANKS1A and CCND1 mRNAs). All five candidates were more downregulated in the trapped excited-state form compared with the wild type, with a 50–80% increase in downregulation efficiency (Fig. 4d, Extended Data Fig. 7 and Methods), suggesting that conformational switching of bulged miRNA–mRNA complexes is a general mechanism for modulating downregulation efficiency.

Discussion

Although seed matching is important, it is only the first step of the RISC cycle. Subsequently, it is thought that nucleation from the 3′-helix can propagate towards the central region. This, together with disengagement of the miRNA 3′-end from the PAZ domain, leads to an active complex, or rather the final step in the RISC activity cycle[1].

We propose that, in the case of miR-34a–mSirt1, this process is mediated by a conformational transition that is triggered by gG8 switching its base-pairing partner. In its ground state, miR-34a–mSirt1 adopts a 7-mer–A1 seed, closed by a weak base pair (g7)—better described as a 6-mer–A1 seed—that is unable to fully displace helix-7 of Ago2 (Fig. 2a). The ground state accesses a distribution of interhelical bend angles that place the miR-34a 3′-end towards the PAZ domain, favouring initial target engagement and nucleation of the 3′-helix[1] (Fig. 4f). During the ground-state to excited-state transition, gG8 repositions to the seed helix and pairs with tU21, resulting in an extended 8-mer–GU seed. The rearrangement of gG8 causes coaxial stacking of the two helices and therefore release of the 3′-end of miR-34a from the PAZ domain, reorienting the RNA duplex towards PIWI domain (Fig. 3e–g) in the simulated structures. This process is accommodated by concerted widening of the N-PAZ channel[12], which facilitates binding of the new stem-to-stem orientation to the cleft and repositioning along the PIWI-N channel in a second binding mode. This excited-state conformation is similar to the catalytically competent state reported for prokaryotic Ago[4,5] (Figs. 3e–g, 4e and Extended Data Fig. 6a, b); moreover, a recent human Ago2 structure confirms that the 3′-helix is mobile[14].

We thus propose that the ground-state to excited-state transition described here provides a mechanism to achieve an active, 'catalytically competent' RISC, promoting mRNA downregulation[1,29]. Although Ago2-bound miRNA is not known to cleave centrally bulged targets, it is possible that these conformational changes enable the RISC to achieve multiple turnovers, which will increase downregulation of the target mRNA[30,31] (Fig. 4f).

Our biophysical and in-cell functional results support this hypothesis, showing a roughly two-fold increase in downregulation upon excited-state stabilization while maintaining RNA–RNA stability. We find that five selected mRNA targets of miR-34a show similar increases in downregulation efficiency when trapped in their excited state (Fig. 4d). Thus the mechanism proposed here could be a widespread feature of bulged binding sites containing partial or extended 3′-pairing.

We have shown that the structural transitions of the guide–target RNA modelled in the RISC provide a mechanistic explanation for bulged complexes, enabling a more accurate prediction of target downregulation by miRNAs. With ever-increasing interest in adapting RNA-guided nuclease machineries for therapeutic, diagnostic and technology applications, we suggest that leveraging the power of RNA conformational dynamics will lead to the design of better guide RNAs, as well as a deeper understanding of these macromolecular complexes.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2336-3.

1. Bartel, D. P. Metazoan microRNAs. *Cell* **173**, 20–51 (2018).
2. Marušič, M., Schlagnitweit, J. & Petzold, K. RNA dynamics by NMR spectroscopy. *ChemBioChem* **20**, 2685–2710 (2019).
3. Ebrahimi, P., Kaur, S., Baronti, L., Petzold, K. & Chen, A. A. A two-dimensional replica-exchange molecular dynamics method for simulating RNA folding using sparse experimental restraints. *Methods* **162-163**, 96–107 (2019).
4. Wang, Y. et al. Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. Nature 461, 754–761 (2009).
5. Sheng, G. et al. Structure-based cleavage mechanism of Thermus thermophilus Argonaute DNA guide strand-mediated DNA target cleavage. *Proc. Natl Acad. Sci. USA* **111**, 652–657 (2014).
6. Elkayam, E. et al. The structure of human argonaute-2 in complex with miR-20a. *Cell* **150**, 100–110 (2012).

7. Schirle, N. T. & MacRae, I. J. The crystal structure of human Argonaute2. *Science* **336**, 1037–1040 (2012) e
8. Nakanishi, K., Weinberg, D. E., Bartel, D. P. & Patel, D. J. Structure of yeast Argonaute with guide RNA. *Nature* **486**, 368–374 (2012).
9. Grimson, A. et al. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol. Cell* **27**, 91–105 (2007).
10. Nielsen, C. B. et al. Determinants of targeting by endogenous and exogenous microRNAs and siRNAs. *RNA* **13**, 1894–1910 (2007).
11. Klum, S. M., Chandradoss, S. D., Schirle, N. T., Joo, C. & MacRae, I. J. Helix-7 in Argonaute2 shapes the microRNA seed region for rapid target recognition. *EMBO J*. **37**, 75–88 (2018).
12. Schirle, N. T., Sheu-Gruttadauria, J. & MacRae, I. J. Structural basis for microRNA targeting. *Science* **346**, 608–613 (2014).
13. Filipowicz, W., Bhattacharyya, S. N. & Sonenberg, N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat. Rev. Genet*. **9**, 102–114 (2008).
14. Sheu-Gruttadauria, J., Xiao, Y., Gebert, L. F. & MacRae, I. J. Beyond the seed: structural basis for supplementary microRNA targeting by human Argonaute2. *EMBO J*. **38**, e101153 (2019).
15. Wee, L. M., Flores-Jasso, C. F., Salomon, W. E. & Zamore, P. D. Argonaute divides its RNA guide into domains with distinct functions and RNA-binding properties. *Cell* **151**, 1055–1067 (2012).
16. He, L. et al. A microRNA component of the p53 tumour suppressor network. *Nature* **447**, 1130–1134 (2007).
17. Yamakuchi, M., Ferlito, M. & Lowenstein, C. J. miR-34a repression of SIRT1 regulates apoptosis. *Proc. Natl Acad. Sci. USA* **105**, 13421–13426 (2008).
18. Agarwal, V., Bell, G. W., Nam, J.-W. & Bartel, D. P. Predicting effective microRNA target sites in mammalian mRNAs. *elife* **4**, e05005 (2015).
19. Vecenie, C. J. & Serra, M. J. Stability of RNA hairpin loops closed by AU base pairs. *Biochemistry* **43**, 11813–11817 (2004).
20. Ulrich, E. L. et al. BioMagResBank. *Nucleic Acids Res*. **36**, D402–D408 (2008).
21. Parisien, M. & Major, F. The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature* **452**, 51–55 (2008).
22. Dethoff, E. A., Petzold, K., Chugh, J., Casiano-Negroni, A. & Al-Hashimi, H. M. Visualizing transient low-populated structures of RNA. *Nature* **491**, 724–728 (2012).
23. Xue, Y. et al. in *Laboratory Methods in Enzymology: RNA* **vol. 558** (eds Woodson, S. A. & Allain, F. H. T.) 39–73 (Academic Press, 2015).
24. Clay, M. C., Ganser, L. R., Merriman, D. K. & Al-Hashimi, H. M. Resolving sugar puckers in RNA excited states exposes slow modes of repuckering dynamics. *Nucleic Acids Res*. **45**, e134 (2017).
25. Broughton, J. P., Lovci, M. T., Huang, J. L., Yeo, G. W. & Pasquinelli, A. E. Pairing beyond the seed supports microRNA targeting specificity. *Mol. Cell* **64**, 320–333 (2016).
26. Brancati, G. & Großhans, H. An interplay of miRNA abundance and target site architecture determines miRNA activity and specificity. *Nucleic Acids Res*. **46**, 3259–3269 (2018).
27. Lingel, A., Simon, B., Izaurralde, E. & Sattler, M. Nucleic acid 3′-end recognition by the Argonaute2 PAZ domain. *Nat. Struct. Mol. Biol*. **11**, 576–577 (2004).
28. Wang, Y., Li, Y., Ma, Z., Yang, W. & Ai, C. Mechanism of microRNA-target interaction: molecular dynamics simulations and thermodynamics analysis. *PLOS Comput. Biol*. **6**, e1000866 (2010).
29. Gebert, L. & MacRae, I. J. Regulation of microRNA function in animals. *Nat. Rev. Mol. Cell Biol*. **20**, 21–37 (2019).
30. Park, J. H., Shin, S.-Y. & Shin, C. Non-canonical targets destabilize microRNAs in human Argonautes. *Nucleic Acids Res*. **45**, 1569–1583 (2017).
31. De, N. et al. Highly complementary target RNAs promote release of guide RNAs from human Argonaute2. *Mol. Cell* **50**, 344–355 (2013).

Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

RNA sample preparation

RNA samples were produced in-house by T7 in vitro transcription[32,33], unless otherwise stated. Modified DNA templates (Integrated DNA Technologies) with oxy-methylated C2′ groups in the first two 5′ nucleotides were used to reduce the 3′-OH heterogeneity of the product[34]. In vitro transcription reactions were supplemented with 20% dimethylsulfoxide (DMSO) to improve reaction yield and to reduce side products[35]. ^{13}C- and ^{15}N-labelled NMR samples were produced by supplementing the transcription reaction with ^{13}C and ^{15}N fully labelled nucleotide triphosphates (Merck Sigma Aldrich). A high-performance liquid chromatography Ultimate3000 uHPLC system (Thermo Scientific) was used to purify the product of interest from abortive transcripts in two chromatographic steps (ion-pair reverse phase and ion-exchange under denaturing conditions) (see Supplementary Methods). hsa-miR-34a-5p 3′-Cy3 labelled and single-stranded mSirt1 in the trapped excited state were purchased from Integrated DNA Technologies as chemically synthesized RNA oligonucleotides purified by RNase-free HPLC purification. A complete list of RNA and DNA sequences used here is given in Supplementary Data S1 Tab 10.

Ago2 preparation and RISC reconstitution

Human Argonaute 2 cloned into the pFastBac HT plasmid was obtained as described[36]. Ago2 was expressed in Sf9 insect cells and purified from the clarified cell lysate through nickel affinity chromatography and gel-filtration chromatography. Sf9 cells were obtained from Invitrogen (catalogue number 11496-015, lot 1296885) and, to our knowledge, were not authenticated. All cell lines were visually inspected throughout the experiments and can be easily identified through their morphology and growth. No misidentified cells were used. The fractions containing Ago2 were pooled together, concentrated and stored at −80 °C. Further details of Ago2 sample preparation are described in the Supplementary Methods.

Purified Ago2 was incubated with a roughly two-fold excess of in vitro transcribed miR-34a in 50 mM Tris-HCl pH 8.0, 300 mM NaCl, 300 mM imidazole and 0.5 mM tris(2-carboxyethyl)phosphine (TCEP) supplemented with 10 μg ml^{-1} bovine serum albumin (BSA; Sigma Aldrich) for 6 h at 37 °C. The assembled RISC (Ago2–miR-34a complex) was then separated from unbound excess RNA by gel filtration chromatography. Loading of the guide miR-34a into the RISC was assessed by an improved northern blot for the detection of small RNA[37,38]. Further details of RISC reconstitution are given in the Supplementary Methods.

Thermal denaturation monitored by UV absorption

Thermal denaturation monitored by UV absorption at 260 nm (A_{260}) was carried out using an Evolution 260 Bio UV-vis spectrophotometer (Thermo Scientific) equipped with a PCCU1 Peltier control and cooling unit (Thermo Scientific). All samples were dissolved in NMR buffer (15 mM sodium phosphate, 25 mM NaCl, 0.1 mM EDTA, pH 6.5). Fitting of the normalized differential melting curves (DMCs; see Supplementary Methods) allowed for estimation of the melting temperature (T_m) and thermodynamic parameters presented in Extended Data Table 1a and Supplementary Data S1 Tab 7.

EMSA

hsa-miR-34a-5p 3′-Cy3 was incubated at a final concentration of 24 nM with increasing amounts of unlabelled single-stranded partner (mSirt1, trapped excited-state mSirt1 or the complementary strand) in NMR buffer (15 mM sodium phosphate, 25 mM NaCl, 0.1 mM EDTA, pH 6.5) to a final volume of 10 µl. The total reaction volumes were mixed with 10 µl of 100% glycerol (Sigma Aldrich) and subsequently loaded into a 10% non-denaturing Tris-bBorate-EDTA (TBE) polyacrylamide gel. Fluorescence signals relative to the free and bound forms of hsa-miR-34a-5p 3′-Cy3 were quantified using ImageJ software[39]. Fitting of the binding curves to a standard binding isotherm (see Supplementary Methods) allowed for estimation of the equilibrium dissociation constants (K_d) presented in Extended Data Table 1b and Supplementary Data S1Tab 8.

Equilibrium filter binding assay

3′-Cy3-labelled target RNAs (mSirt1, trapped excited-state mSirt1 or scrambled control) were incubated at a constant concentration of 0.5 nM with increasing amounts of Ago2–miR-34a complex in target binding buffer[12] (30 mM Tris-HCl pH 8.0, 100 mM potassium acetate, 2 mM magnesium acetate, 2.5 mM TCEP, 0.005% *v*/*v* NP-40 supplemented with 10 µg ml^{-1} yeast transfer RNA (Sigma Aldrich) and 10 µg ml^{-1} BSA (Sigma Aldrich)) to a final volume of 100 µl and incubated for 1 h at 37 °C. After incubation, samples were readily applied to a DHM-48 dot-blot apparatus (Scie-Plas) and filtered through a nitrocellulose membrane (Amersham Protran, GE Healthcare Life Sciences) and a positively charged nylon membrane (Amersham Hybond-N+, GE Healthcare Life Sciences). Fluorescence signals relative to the free (nylon) and protein-bound (nitrocellulose) forms of 3′-Cy3 target RNAs were quantified using ImageJ software[39]. Fitting of the binding curves to a standard binding isotherm (see Supplementary Methods) allowed for estimation of the K_d values presented in Extended Data Table 1c and Supplementary Data S1 Tab 9.

NMR spectroscopy

All NMR assignment and $R_{1\rho}$ relaxation-dispersion experiments were acquired on a Bruker AVANCE III 600 NMR spectrometer operating at 600 MHz for ^{1}H, equipped with a cryogenically cooled QCI probe.

Sequence-specific resonance assignment. These experiments were performed on ^{13}C and ^{15}N fully labelled RNA samples dissolved in 15 mM Na_2HPO_4/NaH_2PO_4, 25 mM NaCl, 0.1 mM EDTA, pH 6.5. Unless otherwise stated, assignment of aromatic ^{13}C2/C5/C6/C8–^{1}H1′/H2/H5/H6/H8, sugar ^{13}C1′–^{1}H1′ and imino ^{15}N1/N3–^{1}H1/H3 resonances was achieved using a standard set of ^{1}H–^{13}C, ^{1}H–^{15}N two-dimensional HSQCs, three-dimensional ^{1}H-^{13}C-^{15}Ns, ^{1}H-^{15}N-^{15}N correlation spectroscopy (COSY) and ^{1}H–^{1}H nuclear Overhauser effect spectroscopy (NOESY) NMR experiments (all acquired using a mixing time of 175 ms) as described[40], recorded at different temperatures (9.0 °C, 22.4 °C and 35.9 °C; Supplementary Figs. 1, 3 and 4). For the miR-34a–mSirt1 duplex, only a reduced set of imino ^{15}N1/N3–^{1}H1/H3 resonances were assigned using ^{1}H–^{15}N two-dimensional HSQCs, HNN COSY and ^{1}H-^{1}H NOESY NMR experiments (Supplementary Fig. 2). Assigned chemical shifts were deposited to the BMRB[20] for hsa-miR-34a-5p (entry 27225), the miR-34a–mSirt1 bulge (entry 27226) and the miR-34a–mSirt1 trapped excited state (entry 27229).

^{1}H, ^{13}C and ^{15}N $R_{1\rho}$ relaxation-dispersion NMR. These experiments were carried out as described[41–44], using ^{13}C and ^{15}N fully labelled (^{13}C and ^{15}N $R_{1\rho}$) or natural-abundance unlabelled (^{1}H $R_{1\rho}$) RNA samples dissolved in 15 mM Na_2HPO_4/NaH_2PO_4, 25 mM NaCl and 0.1 mM EDTA, pH 6.5. In brief, for each spinlock power (ω_{SL}), data points were recorded as a function of different relaxation delays (T_{EX}). For each residue, variable-delay lists were optimized in order to achieve a maximum decay of 1/3 of the starting peak intensity (T_{EX} = 0 s). To account for a reduced loss in peak intensity for large offsets ($\Omega\, 2\pi^{-1}$), we recorded a subset of off-resonance data sets with an extended variable-delay list comprising longer maximal T_{EX} values; we took care that no additional heating occurred. In all data sets, we discarded data points with signal-to-noise ratios (S/N) of less than 20 for ^{1}H and ^{13}C, and of less than 10 for ^{15}N.

Peak intensities were extracted from deconvoluted one-dimensional data sets and plotted as a function of T_{EX}. $R_{1\rho}$ values were obtained from fitting of the data to a mono-exponential decay and error estimates were computed as one standard deviation (s.d.) using a Monte Carlo simulation method[45] with 500 iterations. Potential artefacts (for example, arising from Hartmann–Hahn matching conditions or strong ^{1}H–^{1}H and ^{13}C–^{13}C homonuclear coupling that results in deviation from mono-exponential behaviour) were excluded from subsequent analysis by discarding exponential fits with R^2 values of less than 0.985. $R_{1\rho}$ values as a function of ω_{SL} were subsequently fitted, using the Laguerre approximation[46] (see Supplementary Methods equation (5)), and assuming an absence of exchange ($R_{ex} = 0$), a fast exchange regime (reduced Laguerre form where k_{ex} is much greater than $\Delta\omega$; Supplementary Methods equation (6)), two-state exchange (Supplementary Methods equation (7)) or three-state exchange (Supplementary Methods equation (8) and (9)) using the models and fitting methods further described in the Supplementary Methods. Selection of the best-fitting model for each data set was performed using a statistical *F*-test[23]. Degrees of freedom were calculated as the number of data points (as represented by values in Supplementary Tables 1, 5) minus the number of fitted parameters for each model (two for no exchange, three for reduced Laguerre, five for Laguerre approximation two-state, and eight for Laguerre approximation three-state, for the global fits indicated in Supplementary Table 1). Fitted parameters, reduced χ^2 values resulting from the fit and exact *P*-values from the F-tests (one-sided) for each data set are reported in Supplementary Data S1 Tabs 1, 2, 4.

We carried out global fitting by assuming the presence of one collective exchange process to a minor populated state (ES^{G}), characterized by the global parameters $k_{ex}{}^{G}$ (the global exchange rate) and pop$_{ES}{}^{G}$ (population of ES^{G}) shared across the data sets. Each data set was fitted using the best-fitting model resulting from the individual fits and the fitted parameters as initial guesses for the global fit using a two-state (Supplementary Methods equation (7)) and a three-state (Supplementary Methods equation (8) and (9)) exchange model. For those data sets that were globally fitted using three-state exchange model, we assigned one excited state to the global fit (ES^{G}) while leaving the fitting of the parameters relative to the second state (k_{EX2}, pop$_{ES2}$ and $\Delta\omega_{ES2}$) unconstrained during the fit, fundamentally equivalent to fit them individually. Error estimates of the fitted global parameters were computed as one standard deviation using a Monte Carlo simulation method[45] with 500 iterations. Selection of the best-fitting model was performed using a statistical *F*-test[23], where the simpler fit (a global fit, with a smaller number of parameters) was selected if *P* was greater than 0.05. Degrees of freedoms were calculated as the number of data points minus the number of fitted parameters for each model. Fitted global parameters, reduced χ^2 values resulting from the fit and exact *P*-values from the *F*-test (one-sided) for the global fittings are in Supplementary Data S1 Tab 1.

Exponential fittings, individual and global fittings and model selection were performed using an in-house written Python (2.7) code (https://www.python.org/), available upon request.

Secondary-structure prediction

All secondary-structure predictions were carried out using MC-Fold 1.6.0[21], unless otherwise stated, providing as input the nucleotide sequence of each construct. Structures consisting of two strands where simulated by using a UUCG connection loop.

Chemical-shift distribution of G:C and G:U base pairs

PDB identification codes and nucleotide numbers of guanosines involved in either G:C or G:U base pairs were obtained using RNA FRABASE 2.0[47]. PDB identification codes that have matching BMRB entries were selected using the 'Matched submitted BMRB–PDB entries' list. Subsequently, chemical shifts from ^{1}H1–^{15}N1-assigned couples only were extracted from the BMRB entries, and duplicates and misreferenced couples were removed. A total of 303 G:C and 63 G:U unique ^{1}H1–^{15}N1 couples were obtained (Fig. 1d).

All-atom, explicit solvent molecular dynamics simulations

Atomistic simulations of the miR34a–mSirt1 bulge were initialized using starting structures generated by MC-Fold and MC-Sym[21]. All-atom, explicit solvent molecular dynamics simulations were performed using GROMACS 5.0.7[48] and the modified Chen–Garcia force field for RNA[49], including backbone phosphate modifications[50]. The structure was solvated with 6,664 TIP4P-Ew[51] waters in a 6.1-nm cubic box, and salt conditions of 1 M excess KCl were represented by 161 K^{+} and 134 Cl^{-} ions using activity-coefficient calibrated parameters[52]. In order to enhance exploration of diverse bulge conformations using temperature replica-exchange without inadvertently inducing RNA melting, we assigned five harmonic restraints with a force constant of 500 kJ mol^{-1} nm^{-2} on the middle H-bond of the three initial G:C base-pairs and C14:G19 and G13:C20 (tG25:gC4) in the seed region, which are all observed to be well formed under NMR experimental conditions of 9–35.9 °C. The initial structures were energy minimized and equilibrated at a constant pressure of 1 atm, with random initial velocities drawn from a Boltzmann distribution.

Using REMD, we simulated 24 individual replicas spanning a temperature range of 77–147 °C to evaluate the conformational flexibility of the miR34a–mSirt1 bulge. The exchange rate was 25% with attempted temperature swaps every 1,000 steps (2 ps), which is also how often coordinates were saved. Once equilibrated, production simulations were propagated for roughly 670 ns per replica, a total of 16.08 µs of cumulative simulation time. Structural clustering based on all-heavy-atom r.m.s.d. was accomplished using the algorithm of ref. [53] with 30,000 evenly spaced snapshots taken from the lowest temperature replica (27 °C), using a cut-off of 5.0 Å. The most highly populated cluster, which contains more than 60% of all structures in the 27 °C replica, is the ground-state ensemble (Fig. 2g). We also carried out a separate set of REMD simulations consisting of 25 replicas spanning 25–77 °C, using the same settings as above. Each replica was sampled for 478 ns for a cumulative total of 11.95 µs, and identical cluster analysis was carried out on the 25 °C replica. Details of REMD simulations of the miR-34a bulge excited state and trapped excited state, as well as interhelical bending-angle distributions, are further described in the Supplementary Methods.

Alignment of ground-state/excited-state ensembles into the Ago2 crystal structure. We initially aligned 250 randomly picked snapshots from each REMD ensemble (ground state, excited state and trapped excited state) into the 4W5T PDB structure[12]. Each simulation structure was aligned such that the backbone phosphate positions of bases g2–g8 matched those of the crystal structure. For visual clarity, only 20 of the 250 conformations are graphically depicted in Extended Data Fig. 5.

Slow-growth simulations of insertion of the ground state/excited state into the Ago2 complex. In order to ascertain the ability of the Ago2 protein to physically accommodate the miR-34a–mSirt1 RNA complexes in the ground and excited states, we inserted representative snapshots from each ensemble into the Ago2 protein using slow-growth binding simulations[54]. Starting with the 4W5O PDB structure[12], we deleted the existing partial miRNA–mRNA complex and modelled in missing Ago2 amino acids. The UUCG tetraloop used to anchor the NMR construct was mutated in-place to match the native miR-34a–mSirt1 seed sequence, and the initial RNA conformation was determined by aligning the backbone positions of bases g2–g8 to match the crystallographic RNA seed helix. The RNA was then inserted using a slow-growth process in which RNA–protein van der Waals and electrostatic interactions were completely decoupled at $t = 0$ s, and then linearly increased to 100% interaction in a 100 ps stochastic dynamics simulation at 47 °C, with 1 fs time steps. This method succeeds only if the RNA can be accommodated by flexing of the protein to resolve minor steric overlaps. Successful slow-growth attempts were then solvated in explicit solvent and ions, minimized, and simulated for a roughly 10 ns N,P,T simulation at 25 °C and 1 atm. The conformations shown in Fig. 3e–g are from the final frames of these simulations. The structural models resulting from slow-growth insertion of the ground-state, excited-state and trapped excited-state RNA into the Ago2 protein have been deposited in Model Archive (www.modelarchive.org) under accession codes ma-bc9uo, ma-z54y4 and ma-g8e5z.

Plasmids. All mRNA-targeting DLR[55] plasmids were generated by cloning a synthetic double-stranded DNA (Supplementary Table 10a) into the XhoI and NotI restriction sites of wild-type psiCHECK2-miR-34 (ref. [56]). The fully complementary binding site is the unmodified psiCHECK2-miR-34 WT plasmid[56]. As a negative control, we used the mutated hsa-miR-34a-5p binding site of psiCHECK2-miR-34 MT[56]. These plasmids were a gift from J. Weidhaas (Addgene plasmids 78258 and 78259). The newly generated plasmids were verified by sequencing.

Cell lines and culture. HEK 293T cells were obtained from ATCC (catalogue number CRL-11268) and authenticated by short tandem repeat (STR) analysis by the manufacturer. These cells were used soon after purchase and therefore were not tested for mycoplasma contamination. All cell lines were visually inspected throughout the experiments and can be easily identified by their morphology and growth. No misidentified cells were used. For DLR, HEK 293T cells were cultured in Dulbecco's modified essential medium (DMEM, Gibco) supplemented with 10% fetal calf serum (FCS, Gibco).

DLR assay. HEK 293T cells were seeded 24 h before transfection in 12-well plates. Cells were transfected at 70–90% confluency with 1.6 µg of plasmid DNA and with or without 40 pmol of hsa-miR–34a-5p/hsa-miR–34a-3p (guide/passenger) duplex using lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol. After 24 h, cells were washed with phosphate-buffered saline (PBS) once, and luciferase activity was measured with a DLR assay system (Promega) according to the manufacturer's protocol, using a Promega GloMax 96 microplate luminometer, with a 1-s delay and 10-s integration time. For each sample, the signal corresponding to the *Renilla* luciferase activity was acquired and normalized relative to the firefly luciferase signal. Samples without co-transfected miR-34a were set to 100%, and downregulation of samples co-transfected with miR-34a was calculated on this basis. Results show the average and standard deviation of at least three independent biological replicates. For statistical analysis, we performed unpaired, two-sided single (Fig. 3d) and multiple (Fig. 4d) *t*-tests. Error bars represent one s.d. $^{**}P \leq 0.01$. Details from the fit are presented in Supplementary Data Tab 12.

Predicted target screening of GC-to-GU switches

In total we downloaded 28,653 3′-UTR sequences, including all isoforms, of all 19,432 human protein-coding genes from TargetScan[18]. The sequences were bioinformatically screened for putative mir-34a targets using regular expression. Specifically, sequences were selected that included the reverse complementary sequence of a canonical 6-mer–A1, followed by a U or C as the first nucleotide of the bulge. Thereafter, to allow for a bulge of up to six nucleotides, the sequence was unspecified for positions one to five, and the bulge was closed with a C base-pairing with the G from the miR-34a, leading to this conformational switch model (5′-'C[A,G,U,C]{1,5}[U,C]ACUGCCA'–3′).

Each of the 532 mRNA targets (593 with all isoforms) was screened according to its potential for forming different bulge sizes (from one to five nucleotides) with a G:C or a G:U as the closing base pair. Thereafter, the secondary structure of each mRNA-UUCG–miR34a complex was simulated using MC-Fold 2.32[21]; different mRNA lengths were tested, until a maximum of eight nucleotides was added to an mRNA sequence of 22 nucleotides. Each length was screened to identify examples of ground and excited states similar to those observed for Sirt1, and defined according to the following structural features. A ground state was defined as having: first, a non-base-paired U (position t21 in mSirt1) after the seed, followed by a number of unpaired bases equalling the length of the bulge; second, a GC Watson–Crick base pair closing the bulge, followed by two base pairs, in the 3′-helix (Fig. 4c, cluster 1); and third, a second more stringent cluster (cluster 2) described by two additional Watson–Crick base pairs after the GC closing base pair. An excited state is defined as having a U (position t21 in mSirt1) after the seed pairs with the G in position gG8 (in miR-34a). For obvious structural reasons, in all clusters, we excluded structures in which the miR-34a sequence was folding onto itself or where shortening of the seed was occurring. Sequences were considered only if the ground state and excited state are present for at least three different lengths, and if all the lengths have at least a ground state and an excited state. Of the five targets tested, only CCND1 and ATG9A were previously confirmed as miR-34a targets[57,58].

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

NMR sequence-specific resonance assignments have been deposited in the Biological Magnetic Resonance Data Bank under accesssion numbers 27225 (hsa-miR-34a-5p), 27226 (the miR34a–mSirt1 bulge) and 27229 (the miR34a–mSirt1 trapped excited state). The plasmids used for the DLR assay were a gift from J. Weidhaas (Addgene plasmids 78258 and 78259). All data and code used for data analysis are available upon request. The ensembles of REMD simulations have been deposited in Model Archive (www.modelarchive.org) under accession numbers ma-bc9uo, ma-z54y4 and ma-g8e5z.

32. Beckert, B. & Masquida, B. in *Rna* 29–41 (Springer, 2011).
33. Baronti, L., Karlsson, H., Marušič, M. & Petzold, K. A guide to large-scale RNA sample preparation. *Anal. Bioanal. Chem*. **410**, 3239–3252 (2018).
34. Kao, C., Zheng, M. & Rüdisser, S. A simple and efficient method to reduce nontemplated nucleotide addition at the 3′ terminus of RNAs transcribed by T7 RNA polymerase. *RNA* **5**, 1268–1272 (1999).
35. Helmling, C. et al. Rapid NMR screening of RNA secondary structure and binding. *J. Biomol. NMR* **63**, 67–76 (2015).
36. De, N. & MacRae, I. J. in *Argonaute Proteins* 107–119 (Springer, 2011).
37. Pall, G. S. & Hamilton, A. J. Improved northern blot method for enhanced detection of small RNA. *Nat. Protocols* **3**, 1077–1084 (2008).
38. Rio, D. C. Northern blots for small RNAs and microRNAs. *Cold Spring Harbor Protocols*, https://doi.org/10.1101/pdb.prot080838 (2014).
39. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
40. Fürtig, B., Richter, C., Wöhnert, J. & Schwalbe, H. NMR spectroscopy of RNA. *ChemBioChem* **4**, 936–962 (2003).
41. Schlagnitweit, J., Steiner, E., Karlsson, H. & Petzold, K. Efficient detection of structure and dynamics in unlabeled RNAs: the SELOPE approach. *Chemistry* **24**, 6067–6070 (2018).
42. Hansen, A. L., Nikolova, E. N., Casiano-Negroni, A. & Al-Hashimi, H. M. Extending the range of microsecond-to-millisecond chemical exchange detected in labeled and unlabeled nucleic acids by selective carbon $R_{1\rho}$ NMR spectroscopy. *J. Am. Chem. Soc.* **131**, 3818–3819 (2009).
43. Nikolova, E. N., Gottardo, F. L. & Al-Hashimi, H. M. Probing transient Hoogsteen hydrogen bonds in canonical duplex DNA using NMR relaxation dispersion and single-atom substitution. *J. Am. Chem. Soc.* **134**, 3667–3670 (2012).
44. Steiner, E., Schlagnitweit, J., Lundström, P. & Petzold, K. Capturing excited states in the fast-intermediate exchange limit in biological systems using ^{1}H NMR spectroscopy. *Angew. Chem. Int. Edn* **55**, 15869–15872 (2016).
45. Metropolis, N. & Ulam, S. The Monte Carlo method. *J. Am. Stat. Assoc.* **44**, 335–341 (1949).
46. Palmer, A. G., III & Massi, F. Characterization of the dynamics of biomacromolecules using rotating-frame spin relaxation NMR spectroscopy. *Chem. Rev*. **106**, 1700–1719 (2006).
47. Popenda, M. et al. RNA FRABASE 2.0: an advanced web-accessible database with the capacity to search the three-dimensional fragments within RNA structures. *BMC Bioinformatics* **11**, 231 (2010).
48. Abraham, M. J. et al. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19–25 (2015).
49. Chen, A. A. & García, A. E. High-resolution reversible folding of hyperstable RNA tetraloops using molecular dynamics simulations. *Proc. Natl Acad. Sci. USA* **110**, 16820–16825 (2013).
50. Steinbrecher, T., Latzer, J. & Case, D. A. Revised AMBER parameters for bioorganic phosphates. *J. Chem. Theory Comput*. **8**, 4405–4412 (2012).
51. Horn, H. W. et al. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys*. **120**, 9665–9678 (2004).

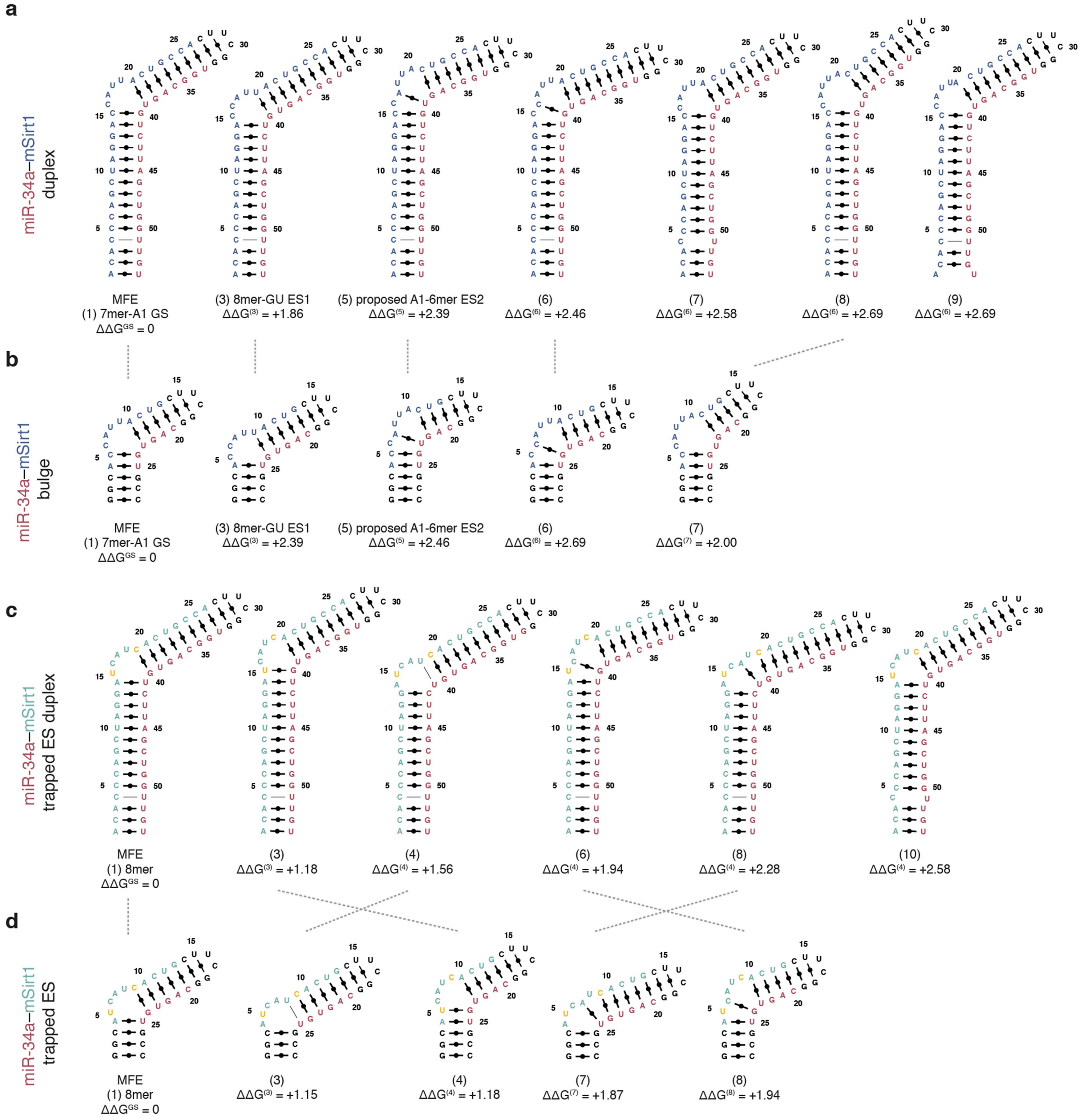
52. Joung, I. S. & Cheatham, T. E., III. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B* **112**, 9020–9041 (2008).
53. Daura, X. et al. Peptide folding: when simulation meets experiment. *Angew. Chem. Int. Edn* **38**, 236–240 (1999).
54. Hu, H., Yun, R. H. & Hermans, J. Reversibility of free energy simulations: slow growth may have a unique advantage (with a note on use of Ewald summation). *Mol. Simul.* **28**, 67–80 (2002).
55. Grentzmann, G., Ingram, J. A., Kelly, P. J., Gesteland, R. F. & Atkins, J. F. A dual-luciferase reporter system for studying recoding signals. *RNA* **4**, 479–486 (1998).
56. Salzman, D. W. et al. miR-34 activity is modulated through 5′-end phosphorylation in response to DNA damage. *Nat. Commun.* **7**, 10954 (2016).
57. Sun, F. et al. Downregulation of CCND1 and CDK6 by miR-34a induces cell cycle arrest. *FEBS Lett.* **582**, 1564–1568 (2008).
58. Huang, J. et al. miR-34a modulates angiotensin II-induced myocardial hypertrophy by direct inhibition of ATG9A expression and autophagic activity. *PLoS One* **9**, e94382 (2014).

Acknowledgements We thank the members of the Petzold and the E. R. Andersson laboratories for insightful discussions; N. H. Hopkins for providing the curated chemical shifts and for contributing to the MC-Fold secondary-structure-prediction script; I. MacRae for the gift of plasmid expressing human Ago2; and the Lund University Protein Production Platform (LP3) (Lund, Sweden) and the Protein Science Facility at the Karolinska Institute Department of Medical Biochemistry and Biophysics (Stockholm, Sweden) for help with human Ago2. K.P. acknowledges funding from the Swedish Research Council (grant number 2014-04303), the Swedish Foundation for Strategic Research (project number ICA14-0023), Harald och Greta Jeansson Stiftelse (JS20140009), Carl Tryggers stiftelse (CTS14-383 and 15-383), Eva och Oscar Ahréns Stiftelse, Åke Wiberg Stiftelse (467080968 and M14-0109), Cancerfonden (CAN 2015/388), the Karolinska Institute Department of Medical Biochemistry and Biophysics (grant number KID 2-3707/2013 and support for the purchase of a 600-MHz Bruker NMR spectrometer) and Ragnar Söderberg Stiftelse (M91/14). A.A.C. acknowledges support from National Science Foundation (grant MCB1651877). B.F. acknowledges funding from the Strategic Research Area (SFO) program of the Swedish Research Council (VR) through Stockholm University. This work used the computational resources of the Extreme Science and Engineering Discovery Environment (XSEDE) (allocation TG-MCB140273 to A.A.C.), which is supported by National Science Foundation grant number ACI-1548562. J.S. acknowledges funding through a Marie Sklodowska-Curie Individual Fellowship (EU H2020/project number 747446).

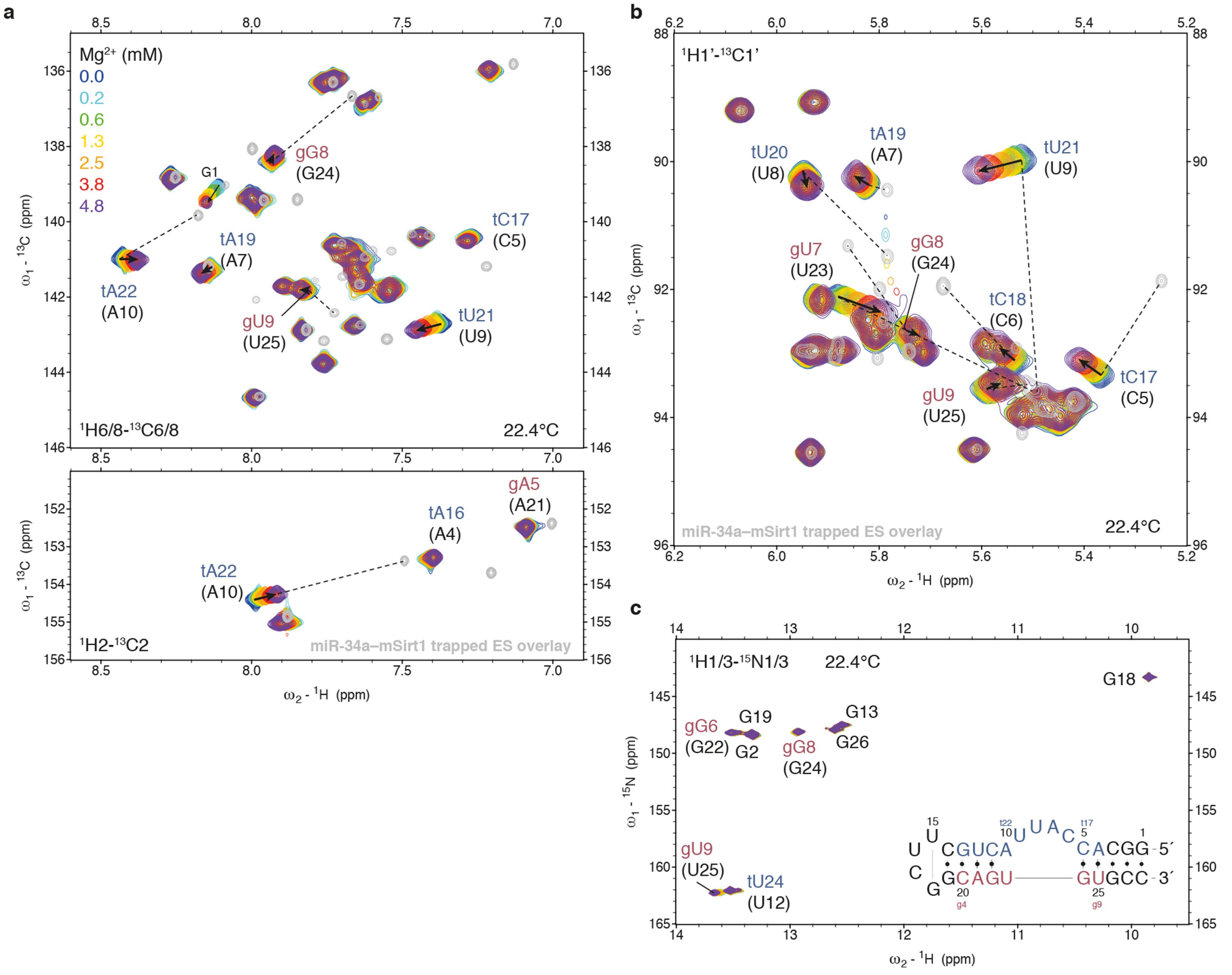
Author contributions L.B. and K.P. conceived the project. L.B. carried out most of the experiments and data analysis, with assistance from I.G., E.S., S.F.S., K.P. and L.S. P.E. and A.A.C. carried out the REMD simulations and slow-growth docking. I.G. and B.F. performed the target screening of GC-to-GU switches. J.S. and C.F. provided advice on design and analysis. A.A.C. supervised computational work and K.P. supervised the overall work. L.B., K.P. and A.A.C. wrote the manuscript and all authors contributed to the final version.

Competing interests The authors declare no competing interests.

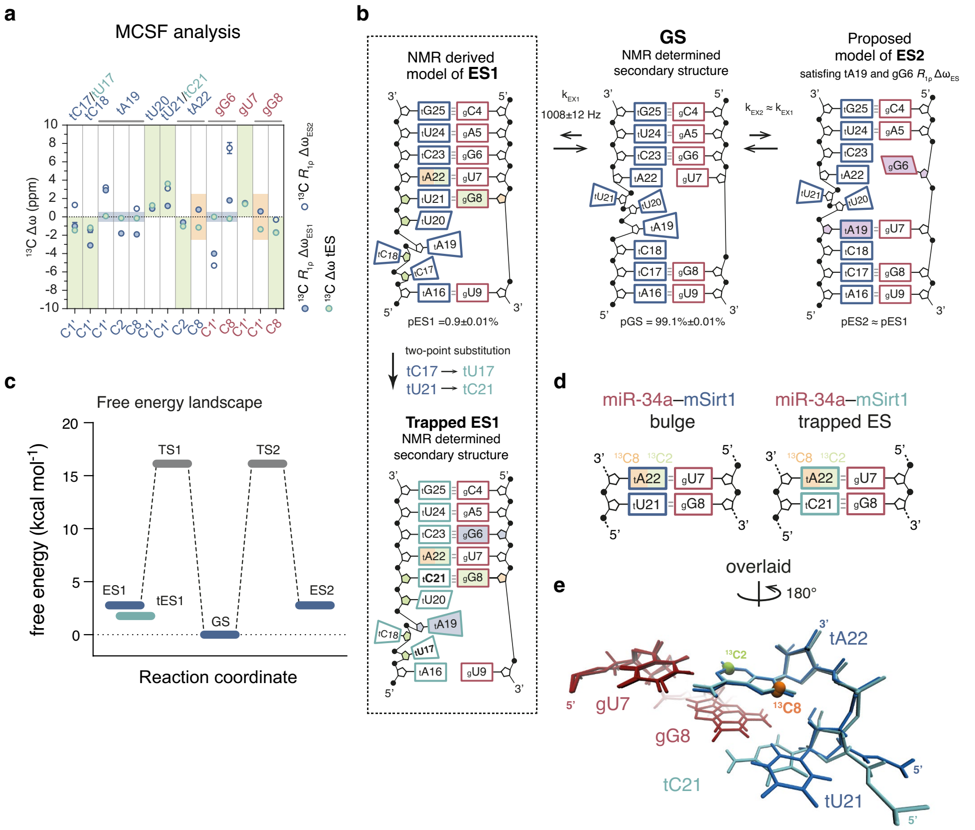
Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2336-3.
Correspondence and requests for materials should be addressed to K.P.
Reprints and permissions information is available at http://www.nature.com/reprints.



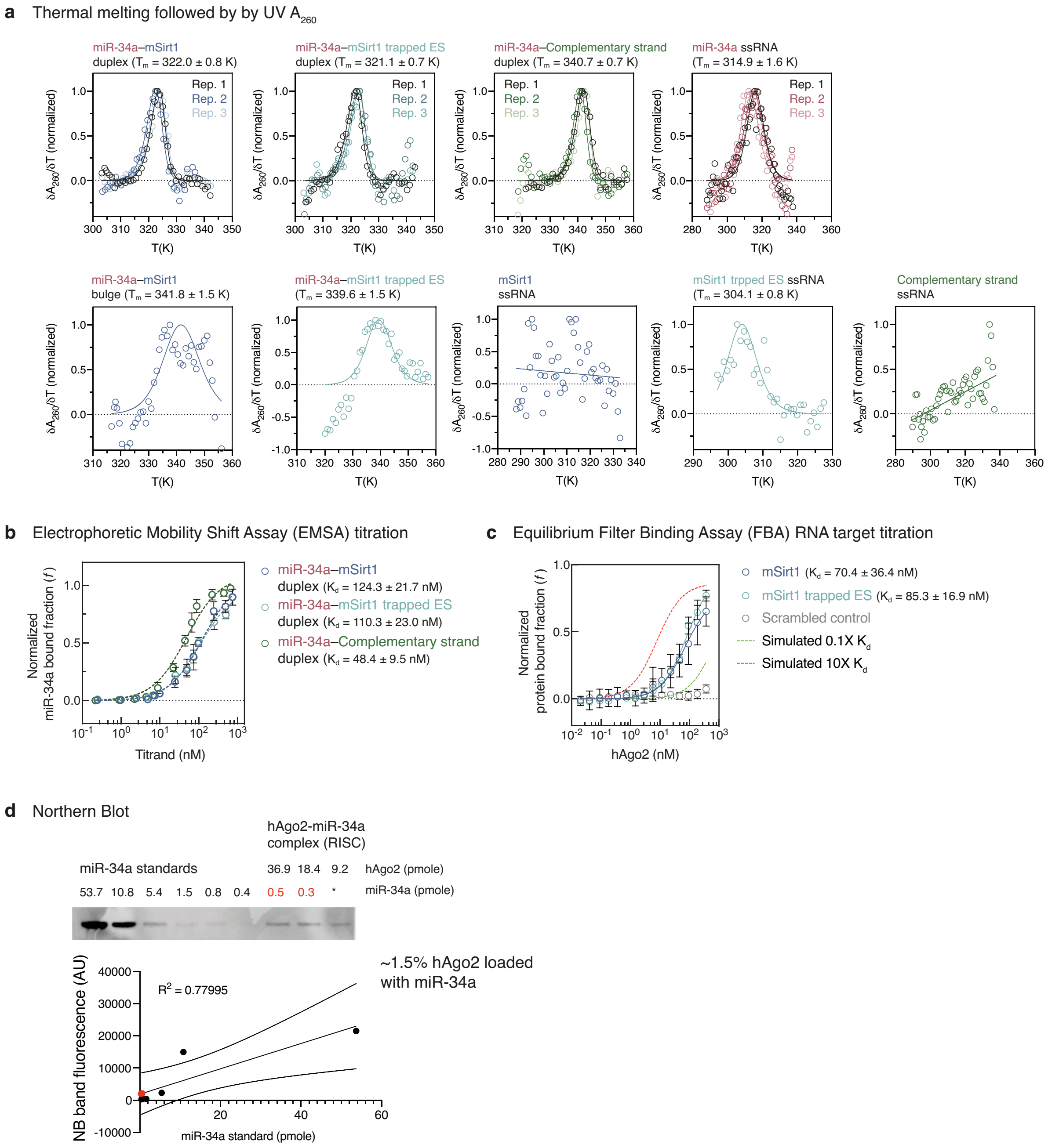
Extended Data Fig. 1 | Secondary-structure prediction using MC-Fold. Secondary-structure rearrangements among the ten lowest-energy structures were calculated using MC-Fold[21]. Ranking (numbers in parentheses) according to the predicted energy difference, based on the minimum free energy (MFE), is indicated in each label (ΔΔG(*n*) in units of unreferenced kcal mol^{-1}, as described[21]). Secondary structures with a single base-pair opening in the cUUCGg region are omitted. **a**, The miR-34a–mSirt1 duplex connected by a cUUCGg loop (black). The MFE corresponds to a 7-mer–A1 binding site. Suboptimal structures (3) and (5) suggest possible modulation of the binding site to a 8-mer–GU and an 6-mer–A1 configuration, respectively. **b**, miR-34a–mSirt1 bulge construct, comprising a cUUCGg loop and a closing stem (black). The secondary-structure distribution of the miR-34a–mSirt1 bulge follows the same trend as the full-length duplex; dashed lines connect identical bulge structures. Suboptimal structures were used to validate or reject models of excited-state (ES) secondary structures on the basis of $R_{1\rho}$ NMR relaxation-dispersion data. Structure (1), with the MFE, corresponds to the assigned ground-state structure (GS). Structure (3) satisfies the ^{1}H$_{1}$ and ^{15}N$_{1}$ $R_{1\rho}$ NMR relaxation-dispersion data on gG6(G24), being G:U base paired with tU20(U9). Structure (5) is mutually exclusive with (3) in structural terms and satisfies the ^{13}C $R_{1\rho}$ NMR relaxation-dispersion data measured on tA19 that indicate this residue adopting a base-paired conformation. Therefore we propose structure (3) as ES1 and structure (5) as ES2. Conformations (6) and (7) do not agree, and partially clash, with our $R_{1\rho}$ NMR relaxation-dispersion data and can therefore be excluded as excited states. **c**, miR-34a–mSirt1 trapped excited-state duplex connected by a cUUCGg loop (black). Substituted nucleotides used to trap the excited state are in yellow. The MFE corresponds to a 8-mer binding site. **d**, miR-34a–mSirt1 (turquoise) trapped excited-state construct comprising an cUUCGg loop and a closing stem (black). Substituted nucleotides used to trap the excited state are in yellow. The secondary-structure distribution of the miR-34a–mSirt1 trapped excited state follows a similar trend as the full-length duplex; dashed lines connect identical bulge structures.



Extended Data Fig. 2 | Mg^{2+} titration of the miR-34a–mSirt1 bulge followed by NMR. Shown are HSQC overlays of different Mg^{2+} titration steps. **a**, ^{1}H–^{13}C aromatic 2/6/8 HSQC. **b**, ^{1}H–^{13}C sugar 1′ HSQC. **c**, ^{1}H–^{15}N imino 1/3 HSQC. The titration steps are colour-coded (**a**, top left). Additional overlay of the miR-34a–mSirt1 trapped ES is shown in grey in **a**, **b**. Arrows indicate the chemical-shift trajectory during titration. Dashed lines connect equivalent peaks in the miR-34a–mSirt1 bulge and trapped ES constructs.

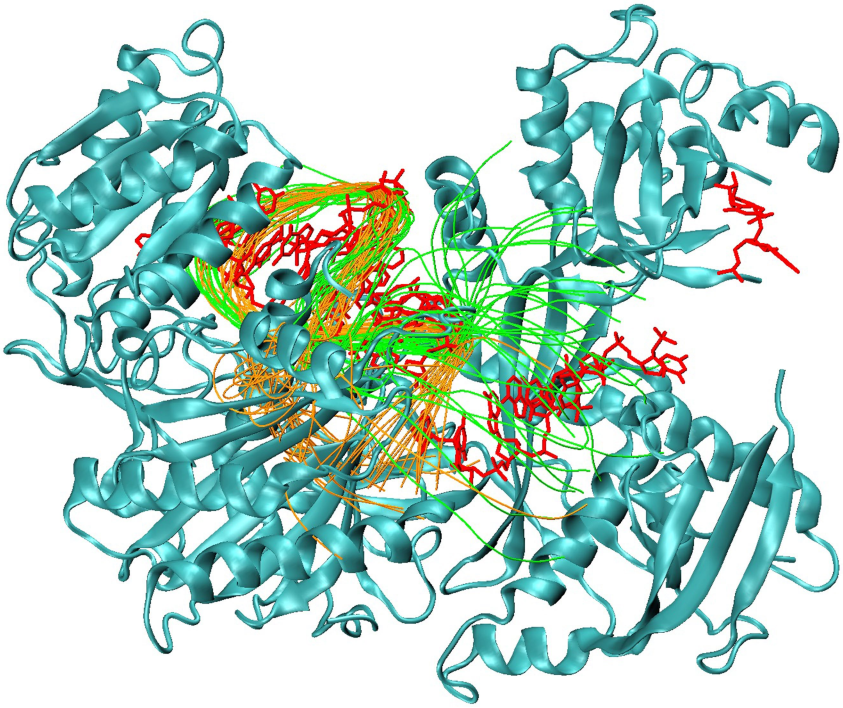


Extended Data Fig. 3 | MCSF analysis of the miR-34a–mSIRT1 bulge and trapped excited state, and analysis of ^{13}C tA22C8 outliers. a, **b**, We used the MCSF approach[22] to cross-validate our candidate excited state (ES1), modelled using $R_{1\rho}$-derived ground-state-to-excited-state chemical-shift differences (**a**, ^{13}C $R_{1\rho}$ $\Delta\omega$ data, blue dots; **b**, left). p_{ES} refers to the excited-state population (pop_{ES} in the main text). We also generated an ES1 mimic (trapped ES1) using a two-point substitution, predicted to stabilize the proposed conformation (**b**, bottom). For each reporter atom, we compared ^{13}C $R_{1\rho}$ $\Delta\omega$ data with the chemical-shift differences derived from the assignment of the bulge and the trapped ES constructs (**a**, ^{13}C $\Delta\omega$ trapped ES (tES) data, turquoise dots). In **a**, The MCSF analysis validates our ES1 model (green shading), with exceptions arising from the limitations of the mimic (orange shading) and from the presence (violet shading) of a second ES (ES2, **b**, right). Errors for $R_{1\rho}$ relaxation-dispersion-derived $\Delta\omega$ represent 1 s.d. from fitting (see also Supplementary Methods). In **b**, the proposed model for ES2 satisfies the ^{13}C $R_{1\rho}$ $\Delta\omega$ data measured for tA19 and gG6. GS, ground state. **c**, The free-energy landscape for the entire star-like three-state exchange process. (The MCSF analysis and ES2 are discussed further in the Supplementary Information, Discussion 5.) The transition coeffient (κ), is assumed to be 1 (ref. [23]), so the transition-state energies (TS1 and TS2), calculated using Supplementary equation (11), must be considered an upper limit of this exchange process. **d**, **e**, The substitution site (tU21 to tC21) perturbs the chemical environment of tA22C8 that is directly neighbouring the substituted nucleobase (orange sphere in **e**). Conversely, tA22C2 (green sphere), pointing towards the miR-34a strand (red), experiences an equivalent chemical environment in the bulge (blue) and trapped ES (turquoise) constructs. This explains the inconsistency in the MCSF profile for tA22C8 (Supplementary Fig. 12a, orange box). **d**, Secondary structure environment of tA22 in the miR-34a–mSirt1 bulge excited state (left) and trapped ES (right) constructs. The substitution site (tU21 to tC21) is highlighted. **e**, Overlay of average structures of the bulge ES (blue) and trapped ES (turquoise) from REMD ensembles, aligned according to residues gU7 and tA22. Residues gU7, gG8, tU21 and tA22 are shown. tA22C8 and tA22C2 ^{13}C atoms are in orange and green respectively.

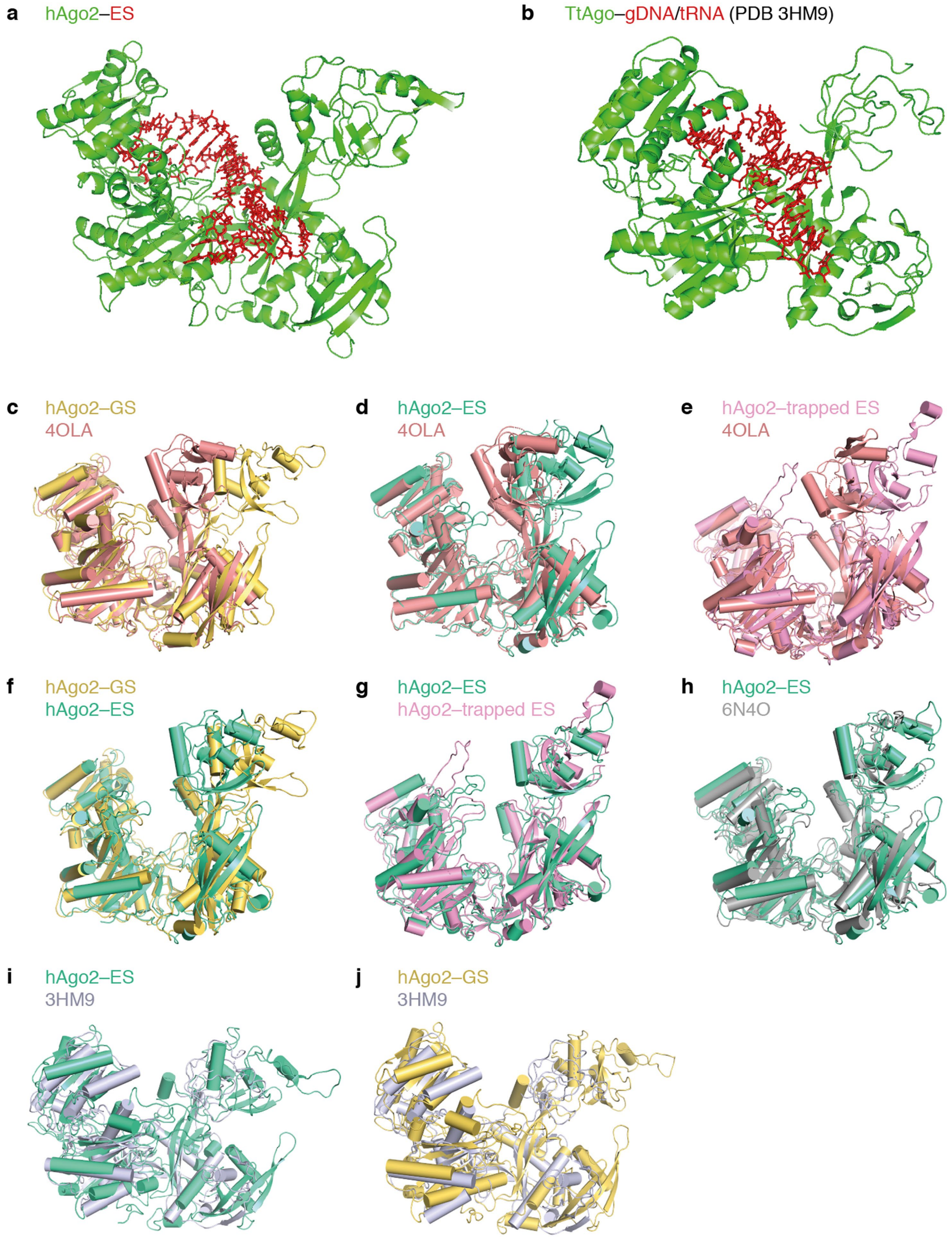


Extended Data Fig. 4 | See next page for caption.

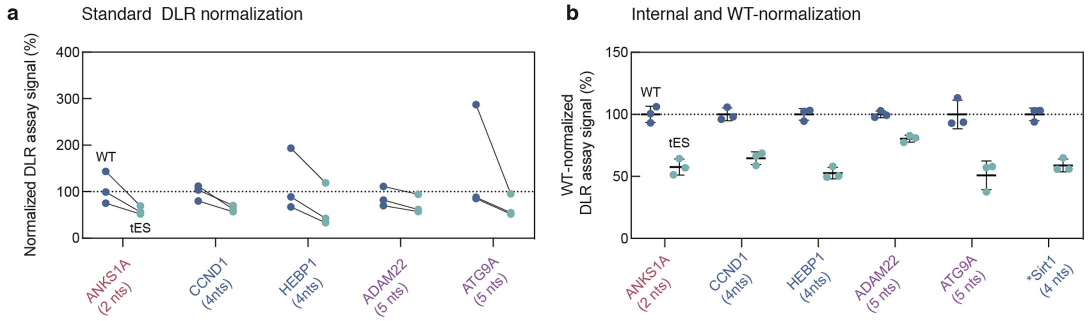
Extended Data Fig. 4 | Biophysical and biochemical characterization of the constructs. a, Individual A_{260} UV melting profiles for the constructs used here. The miR-34a–mSirt1 duplex, miR-34a–mSirt1 trapped ES duplex, miR-34a–complementary-strand duplex and miR-34a single-stranded RNA (ssRNA) were each measured as three technical independent replicates (shown in different colours; $n = 1$). Individual technical replicates are plotted. T_m values are shown as means ± s.d. of fitted T_m values in individual technical replicates ($n = 3$). The other ssRNAs (bottom row) were measured and plotted as individual technical replicates; fitted T_m values are shown with associated confidence intervals of 95% ($n = 1$) as an estimate of the experimental error. Normalized differential melting curves ($\delta A_{260}/\delta T$) are plotted as a function of temperature (in K) (circles) and fitted to Supplementary equation (1a) or (1b) (curves), depending on the molecularity of the system. **b**, EMSA titration profiles for the miR-34a–mSirt1 duplex, miR-34a–mSirt1 trapped ES duplex and miR-34a–complementary-strand duplex, measured as three independent technical replicates. The ratio of bound to total miR-34a 3′-Cy3 is plotted as a function of titrand concentration (circles) and fits a standard binding isotherm (line) (Supplementary equation (2)). The plot centre is the mean; error bars represent 1 s.d. from the three independent replicates. Fitted K_d values along with confidence intervals of 95% are shown as an estimate of the experimental error ($n = 3$). Gel images were acquired by detection of Cy3 fluorescence. During the titration, miR-34a 3′-Cy3 was kept at a constant concentration of 24 nM, setting the sensitivity limit for estimating K_d (Supplementary Fig. 1a–c). mSirt1 and its trapped-ES counterpart are equivalent in their ability to form a stable RNA–RNA duplex with miR-34a. Tighter binding is observed for the complementary strand (48.4 ± 9.5 nM) than for the mSirt1 (124.3 ± 21.7 nM) and trapped-ES mSirt1 (110.3 ± 23.0 nM), providing a control for the dynamic range of K_d estimation. **c**, Equilibrium FBA profiles for mSirt1, mSirt1 trapped ES and a scrambled control, binding to miR-34a-loaded Ago2. The three targets were each measured as three independent replicates and fitted to a standard binding isotherm (line) (Supplementary equation (2)). The plot centre is the mean; error bars represent 1 s.d. from three independent replicates. Fitted K_d values are shown with confidence intervals of 95% (an estimate of the experimental error). As in **c**, mSirt1 and mSirt1 trapped ES are equivalent in their ability to form a stable ternary complex within RISC. The simulated data set (dotted lines) indicate curves corresponding to K_d values ten times lower (red) or ten times higher (green) than the average value for mSirt1 and mSirt1 trapped ES, providing a frame for the amplitude of our experimental error. **d**, Top, northern blot showing the detection of miR-34a loaded in Ago2. Bottom, a standard calibration curve (using naked miR-34a), used to obtain an estimate of miR-34a in RISC. The centre calibration curve was used to calculate R^2. The two outer curves indicate the 95% confidence interval of the calibration-line fit (from a single repeated experiment). The average ratio of Ago2 and miR-34a-loaded Ago2 (both in pmole) was used to obtain the fraction of Ago2 loaded with our guide (roughly 1.5%). The complete lists of fitted parameters for UV melting, EMSA titration, FBA titration and northern blot are in Supplementary Table 1a–d. The complete fitting analyses of UV melting, EMSA titration and FBA titration are in Supplementary Tables 7–9.



Extended Data Fig. 5 | Crystal structure of Ago2 overlaid with REMD ensembles. Superposition of ground state (green) and excited state (orange) conformational ensembles on the Ago2 crystal structure (PDB code 4W5T), with seed sequences aligned to crystallographic miRNA–mRNA positions (red). Although the seed orientations are comparable, the ground-state and excited-state conformations sample different space within Ago2.



Extended Data Fig. 6 | Slow-growth insertion of excited-state RNA into Ago2 predicts the ability of bulged miRNA–mRNA complexes to access an alternative dsRNA-binding mode of Ago2. Slow-growth induced-fit Ago2 structures are compared with existing X-ray structures (whose PDB identification codes are shown in the figure) via structural alignment. **a**, Ago2 after induced fit with ES RNA binds in the PIWI-adjacent groove rather than the PAZ domain. **b**, The *Thermus thermophilus* (Tt) Ago crystal structure similarly shows DNA/RNA-duplex binding in the analogous PIWI-adjacent groove. **c**–**j**, The root mean square deviation (r.m.s.d.) for each indicated pair of Ago structures was measured after structural alignment either of all protein atoms, or excluding the PAZ domain, PIWI loops and helix-7 atoms ('subset aligned'; these excluded atoms still count towards the r.m.s.d.). The subset-aligned structures show that most of the r.m.s.d. difference arises from pivoting motions of the PAZ domain, coupled with small shifts in helix-7 and PIWI loops to accommodate the inserted ES RNA structures. **c**, Comparison of slow-growth human Ago2–GS and the existing Ago2 structure (PDB code 4OLA; r.m.s.d. = 2.065 Å (all) and 2.62 Å (subset aligned)). **d**, Comparison of slow-growth Ago2–ES and the 4OLA structure (r.m.s.d. = 1.4 Å (all) and 1.65 Å (subset aligned)). **e**, Comparison of the slow-growth Ago2–trapped ES and the 4OLA structure (r.m.s.d. = 1.9 Å (all) and 2.18 Å (subset aligned)). **f**, Comparison of the slow-growth Ago2–GS with Ago2–ES (r.m.s.d. = 2.1 Å (all) and 2.2 Å (subset aligned)). **g**, Comparison of the slow-growth Ago2–ES with Ago2–trapped ES (r.m.s.d. = 1.6 Å (all) and 1.33 Å (subset aligned)). **h**, Comparison of the slow-growth Ago2–GS and the 6N4O structure (r.m.s.d. = 2.05 Å (all) and 2.065 Å (subset aligned)). **i**, Comparison of the slow-growth Ago2–ES (green) with the 3HM9 structure (r.m.s.d. = 4.52 Å (all)). **j**, Comparison of the slow-growth Ago2–GS with the 3HM9 structure (r.m.s.d. = 3.85Å (all)).



Extended Data Fig. 7 | DLR assay of additional miR-34a targets. We studied five targets of different bulge sizes (see Methods). Individual replicates are plotted as circles; the centre line is the mean; error bars represent 1 s.d from three independent replicates; nts, nucleotides. **a**, Standard DLR normalization (relative to the control condition with no miR-34a duplex transfected). Despite the large variability between replicates, a consistent increase in downregulation (connecting lines) is observed for wild-type (WT) and trapped excited-state (tES) constructs. **b**, When the data sets are internally normalized and the WT condition is set to 100% (mean value), the variation due to experimental replicas is attenuated and the trend observed in **a** is maintained.

Extended Data Table 1 | T_m and K_d fitted parameters

a UV A_{260} thermal melting	mean T_m (K)	T_m (K)	h	ΔH (kJ mol^{-1})	ΔS (kJ K^{-1} mol^{-1})	R^2
miR-34a–mSirt1 duplex	322.0 ± 0.8	322.5 ± 0.2	-280.0 ± 17.9	-750.8 ± 398.0	-2.3 ± 1.2	0.9657
		321.1 ± 0.3	-257.4 ± 24.2	-687.2 ± 537.2	-2.1 ± 1.7	0.9339
		322.5 ± 0.3	-254.3 ± 19.3	-681.9 ± 430.3	-2.1 ± 1.3	0.9513
miR-34a–mSirt1 trapped ES duplex	321.1 ± 0.7	320.3 ± 0.3	-259.2 ± 24.9	-690.3 ± 550.3	-2.2 ± 1.7	0.9311
		321.3 ± 0.7	-249.9 ± 43.2	-667.6 ± 958.5	-2.1 ± 3.0	0.7994
		321.7 ± 0.4	-277.4 ± 29.7	-742.0 ± 660.6	-2.3 ± 2.1	0.9026
miR-34a-Complentary strand duplex	340.7 ± 0.7	340.9 ± 0.3	-285.6 ± 19.4	-809.5 ± 457.2	-2.4 ± 1.3	0.9657
		339.9 ± 0.5	-321.1 ± 53.9	-907.5 ± 1265.4	-2.7 ± 3.7	0.8243
		341.3 ± 0.3	-349.8 ± 44.8	-992.6 ± 1057.0	-2.9 ± 3.1	0.9023
miR-34a ssRNA	314.9 ± 1.6	316.3 ± 0.8	-97.5 ± 0.8	-256.4 ± 17.3	-0.8 ± 0.1	0.8369
		315.3 ± 0.8	-89.5 ± 0.8	-234.7 ± 19.2	-0.7 ± 0.1	0.8246
		313.1 ± 0.8	-93.5 ± 0.8	-243.5 ± 18.1	-0.8 ± 0.1	0.8388
miR-34a-mSIRT1 bulge	-	341.8 ± 1.5	-77.5 ± 1.6	-220.3 ± 38.7	-0.6 ± 0.1	0.6198
miR-34a-mSIRT1 trapped ES	-	339.6 ± 1.5	-110.1 ± 1.5	-310.9 ± 37.0	-0.9 ± 0.1	0.6477
mSirt1 ssRNA	-	-				-
mSirt1 trapped ES ssRNA		304.1 ± 0.8	-107.7 ± 1.5	-272.3 ± 363.7	-0.9 ± 1.2	0.7834
Complementary strand ssRNA	-	-				-

b EMSA	K_d (nM)	R^2
miR-34a–mSirt1 duplex	124.3 ± 21.7	0.9821
miR-34a–mSirt1 trapped ES duplex	110.3 ± 23.0	0.9732
miR-34a-Complentary strand duplex	48.4 ± 9.5	0.9755

c FBA	K_d (nM)	R^2
miR-34a–mSirt1 duplex	70.4 ± 36.4	0.8791
miR-34a–mSirt1 trapped ES duplex	85.3 ± 16.9	0.9822
Scrambled control	-	-

d Northern blot	%
Estimated hAgo2 loaded with miR-34a	~1.5

a, Thermal denaturation was monitored by UV absorption. Mean ± s.d. T_m values were obtained from three independent replicates. Also shown are parameters derived from fitting of Supplementary equation (1a) or (1b) (Supplementary Methods). Fitting parameters T_m and h are presented with confidence intervals of 95% (as estimates of the experimental error) (T_m is the melting temperature and $h = \Delta H/RT_m$ ($R = 8.31447$ J K^{-1} mol^{-1})). Complete fitting details and statistics are in Supplementary Table 7. **b**, **c**, EMSA and FBA. Parameters derived from fitting of Supplementary equation (2) (see Supplementary Methods). K_d values obtained from the fit are presented with confidence intervals of 95% as estimates of the experimental error ($n = 3$). Complete fitting details and statistics are in Supplementary Tables 8, 9. **d**, Northern blot. The fraction of Ago2 loaded with the guide RNA of interest was estimated by northern blotting (see Supplementary Methods).

nature research

Corresponding author(s): Katja Petzold
Last updated by author(s): Feb 10, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	All software and code details are described in the Method and Supplementary Methods sections.
Data analysis	All software and code details are described in the Method and Supplementary Methods sections.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

NMR sequence specific resonance assignment of hsa-miR-34a-5p (entry 27225), miR34a–mSirt1 bulge (entry 27226) and miR34a–mSirt1 trapped ES (entry 27229) constructs were deposited in the BMRB. The plasmids used in this work were a gift from Joanne Weidhaas (Addgene plasmid #78258 and #78259). hAgo2 expressing plasmid was a gift from (Prof. Ian MacRae, Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA, USA). The structural models resulting from slow-growth insertion of the GS, ES, and trapped ES RNA into the hAGO2 protein have been deposited in Model Archive (www.modelarchive.org) as ma-bc9uo, ma-z54y4, and ma-g8e5z.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size calculation was performed. Cell-based assays were performed in at least 3 independent, biological replicas.
Data exclusions	No data was excluded from analysis.
Replication	Experiments were successfully replicated to ensure that they stably support our findings. The numbers are given for relevant experiments. Cell-based assays were performed at least in 3 independent, biological replicates. Technical replicates were performed where necessary and are stated.
Randomization	Samples were not randomized for analysis.
Blinding	No blinding was performed in this study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study	
☒	☐	Antibodies
☐	☒	Eukaryotic cell lines
☒	☐	Palaeontology
☒	☐	Animals and other organisms
☒	☐	Human research participants
☒	☐	Clinical data

Methods

n/a	Involved in the study	
☒	☐	ChIP-seq
☒	☐	Flow cytometry
☒	☐	MRI-based neuroimaging

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	HEK293T cells were obtained from ATCC (CRL-11268), Sf9 cells were obtained from Invitrogen (Cat no 11496-015, Lot 1296885).
Authentication	All human cells lines from ATCC are authenticated by STR analysis. Sf9 cells were to our knowledge not authenticated. All cell lines were visually inspected throughout the experiments and can be easily identified through their morphology and growth.
Mycoplasma contamination	HEK293T cells (CRL-11268) were readily used after purchase and therefore were not tested for Mycoplasma contamination
Commonly misidentified lines (See ICLAC register)	No misidentified cell lines were used

Electromechanical coupling in the hyperpolarization-activated K$^+$ channel KAT1

https://doi.org/10.1038/s41586-020-2335-4

Received: 23 April 2019

Accepted: 16 March 2020

Michael David Clark[1], **Gustavo F. Contreras**[1], **Rong Shen**[1] & **Eduardo Perozo**[1]✉

Voltage-gated potassium (K$_v$) channels coordinate electrical signalling and control cell volume by gating in response to membrane depolarization or hyperpolarization. However, although voltage-sensing domains transduce transmembrane electric field changes by a common mechanism involving the outward or inward translocation of gating charges[1–3], the general determinants of channel gating polarity remain poorly understood[4]. Here we suggest a molecular mechanism for electromechanical coupling and gating polarity in non-domain-swapped K$_v$ channels on the basis of the cryo-electron microscopy structure of KAT1, the hyperpolarization-activated K$_v$ channel from *Arabidopsis thaliana*. KAT1 displays a depolarized voltage sensor, which interacts with a closed pore domain directly via two interfaces and indirectly via an intercalated phospholipid. Functional evaluation of KAT1 structure-guided mutants at the sensor–pore interfaces suggests a mechanism in which direct interaction between the sensor and the C-linker hairpin in the adjacent pore subunit is the primary determinant of gating polarity. We suggest that an inward motion of the S4 sensor helix of approximately 5–7 Å can underlie a direct-coupling mechanism, driving a conformational reorientation of the C-linker and ultimately opening the activation gate formed by the S6 intracellular bundle. This direct-coupling mechanism contrasts with allosteric mechanisms proposed for hyperpolarization-activated cyclic nucleotide-gated channels[5], and may represent an unexpected link between depolarization- and hyperpolarization-activated channels.

Voltage-gated ion channels couple electric field-driven conformational changes in their voltage-sensing domains (VSDs) to mechanical opening and closing of their pore domains[6–8]. This process of electromechanical coupling underlies the function of both depolarization- and hyperpolarization-activated channels (Extended Data Fig. 1). To better understand the molecular basis of how electromechanical coupling might lead to these two distinct gating polarities, we determined the cryo-electron microscopy (cryo-EM) structure of the hyperpolarization-activated potassium channel KAT1 from *A. thaliana*, and probed the interactions between its voltage-sensing and pore domains using mutagenesis, electrophysiology and modelling. KAT1 is a founding member of the plant inwardly rectifying, potassium-selective ion channel subfamily. Physiologically, these channels tune osmotic potential to hydraulically control stomatal opening in flowering plants[9]. A fully functional construct (KAT1em) (Fig. 1a) spanning the transmembrane region and pseudo cyclic nucleotide-binding domain (CNBD) (Extended Data Fig. 2) was purified and imaged in the gentle detergent digitonin. In the cryo-EM images, KAT1em assembles as a dimer of two tetrameric channels stacking via their cytoplasmic domains (Fig. 1b); although the physiological importance, if any, of this stacking is currently unknown. Focused refinement of the tetramer improved map quality (from a nominal resolution of 3.8 Å for the full dimer of tetramers to a nominal 3.5 Å for the tetramer transmembrane region) and facilitated de novo model building (Extended Data Figs. 2, 3, Extended Data Table 1). KAT1em shares the topology of the CNBD-containing channel family, including a non-domain-swapped subunit arrangement of its voltage-sensing and pore domains, followed by a C-linker and pseudo-CNBD (Fig. 1c, d).

The pore domain of KAT1 displays a closed inner gate, with its narrowest constriction formed by the hydrophobic side chains of I292, whereas the selectivity filter is in a conductive conformation (Fig. 2a, Extended Data Fig. 4). Functionally, these characteristics correspond to the expected closed state at 0 mV (refs. [10–12]). To evaluate the energetics of pore opening, we conducted a local alanine scan of the inner gate region. Six mutants did not produce currents, and seven mutants displayed a range of energetic effects (Fig. 2a, b). On one side of the pore-lining helix S6, L287A—which packs against the S5 helix—promotes channel opening. S5–S6 packing interactions have previously been proposed to stabilize the closed state of the hyperpolarization-activated cyclic nucleotide-gated channel HCN[13] and the potential reduction in van der Waals interactions at this position might facilitate gate opening. By contrast, V299A (at the intracellular end of the helical-bundle gate, nestled against the neighbouring S6) and T288A (towards the middle of S6) promote channel closure. Together, these results suggest the reorganization of S5–S6 and S6–S6 interhelical packing upon channel activation–deactivation.

KAT1em VSDs are arranged as four-helix bundles, each with a centrally located hydrophobic gasket (or plug) formed by the side chains of F102 and V70 (Fig. 2c). There are six arginines on S4, labelled R0–R6 from the extracellular to intracellular end of the helix. R0 (R165),

---

[1]Department of Biochemistry and Molecular Biology, The University of Chicago, Chicago, IL, USA. ✉e-mail: eperozo@uchicago.edu

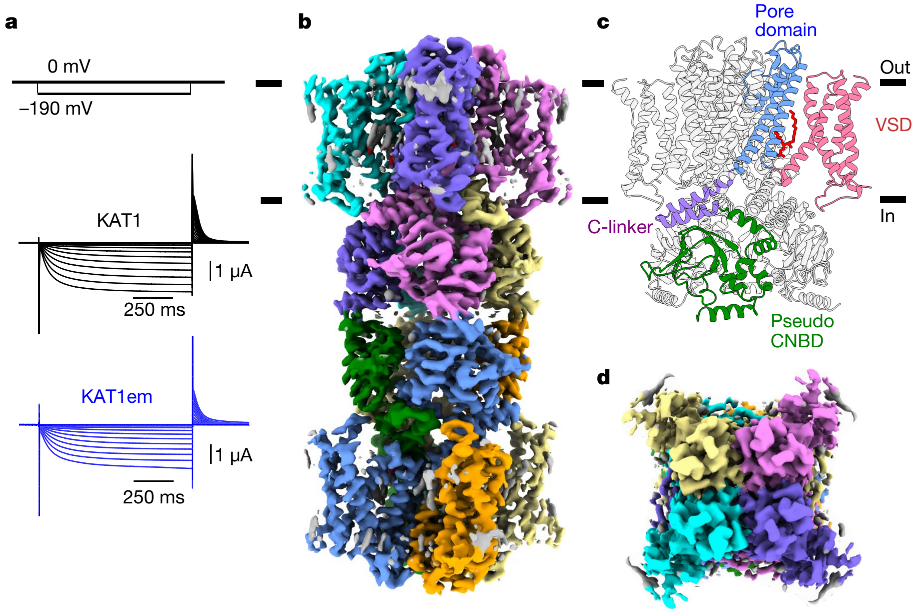


Fig. 1 | Function and architecture of *A. thaliana* KAT1em. a, Macroscopic currents of full-length KAT1 and KAT1em, recorded in *Xenopus* oocytes using a family of hyperpolarizing pulses (top). Data are representative of 11 and 6 biologically independent cells for full-length KAT1 and KAT1em, respectively. **b**, Sharpened cryo-EM density map of the channel octamer, side view. **c**, Ribbon model of KAT1em, with domains labelled. Phospholipid is shown in red stick representation. **d**, Sharpened cryo-EM density map of the channel tetramer, top view (view from extracellular side).

R1 (R171) and R2 (R174) are positioned above the gasket, whereas R3 (R176), R4 (R177) and R5 (R184) are located below the gasket (Fig. 2c, Extended Data Fig. 5a). This conformation of the KAT1 VSD corresponds to an 'up', or depolarized state, which—in the nominal absence of a field (0 mV)—is coupled to a closed pore domain. Limiting-slope analysis in KAT1 has suggested an effective z_e of about 3 *e* per channel (approximately 0.75 *e* per sensor)[3] consistent with R2, and possibly R1, serving as the main sensing charges. Accordingly, mutant channels R174Q and R171Q did not yield currents (data not shown). As described below, mutant cycle and metal-bridge data also indicate that the VSD structure corresponds to an up state and point to a number of residue–residue pairs with interactions that probably change on transitioning to the 'down' state during membrane hyperpolarization.

KAT1em voltage-sensing and pore domains interact through two major interfaces: the first near the intracellular face of the channel (Fig. 3a, b) with the participation of S4 and S5 overlaying the C-linker of the adjacent subunit, and the second near the extracellular side formed by the intercalation of S1 between S4 and S5 of the same subunit (Fig. 3c). At the first, intracellular interface, the extended length of the KAT1em S4 mediates interactions between S4, S5 and the C-linker. The intracellular ends of the S4 and S5 helices come to rest on top of the C-linker, forming a tightly packed interface. Notably, R310 from the

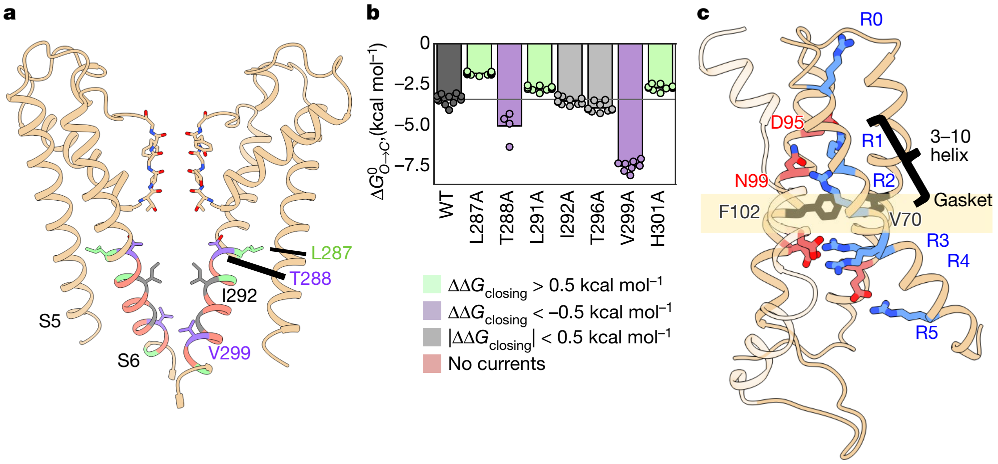


Fig. 2 | KAT1 pore and voltage-sensing domain structure and alanine scanning of pore inner gate region. a, View of the pore, with only two subunits shown for clarity. Sticks are shown for selectivity-filter residues, inner-gate-forming residue I292 and functionally important residues L287, T288 and V299. Residues are coloured by effect of alanine mutagenesis (see inset legend). **b**, Deactivation energies ($\Delta G_{o\to c}$) of alanine mutants, calculated from conductance–voltage ($G(V)$) relations (Extended Data Fig. 4b). $\Delta\Delta G_{closing}$ refers to the difference in deactivation energy between a given mutant and wild type. The wild type (WT; $n = 11$) and L287A ($n = 19$), T288A ($n = 4$), L291A ($n = 10$), I292A ($n = 10$), T296A ($n = 10$), V299A ($n = 8$) and H301A ($n = 10$) mutants are shown; *n* is the number of biologically independent cells. Six mutants did not yield currents: Y290A, N294A, M295A, N297A, L298A, and V300A (data not shown) **c**, Rotated views of the KAT1em VSD. Stick side chains are shown for the hydrophobic gasket: F102 and V70, for key residues on S4: R165 (R0), R171 (R1), R174 (R2), R176 (R3), R177 (R4) and R184 (R5), and for counter-charges or dipoles: E63, D95, N99, D105 and D141.

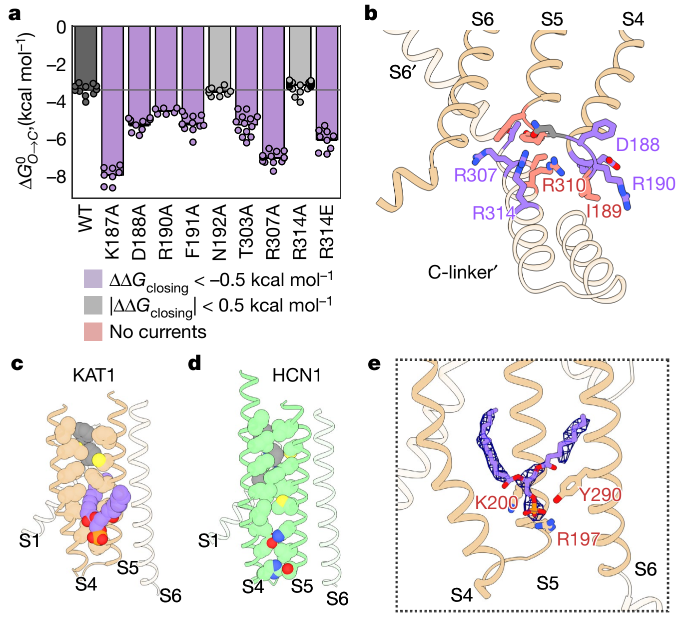


Fig. 3 | The KAT1 VSD–pore interface and lipid-binding conformation. **a**, Deactivation energies of S4–S5 C-linker interfacial mutants calculated from *G*(*V*) relations (Extended Data Fig. 7a). The wild type (*n* = 11) and K187A (*n* = 8), D188A (*n* = 9), R190A (*n* = 6), F191A (*n* = 12), N192A (9), T303A (*n* = 13), R307A (*n* = 14), R314A (*n* = 31) and R314E (*n* = 9) mutants are shown; *n* is the number of biologically independent cells. **b**, Mapping of electrophysiology data from **a**, coloured by the effect of mutation as indicated in legend inset. Key S4–S5 linker residues K187, D188, I189, R190, N192, Y193 and F194, and key neighbouring-subunit C-linker residues T306, R307, R310 and R314 are shown as sticks. **c**, KAT1 upper VSD–pore interface (S1, S4 and S5) residue packing shown as spheres. Bound phospholipid in the hydrophobic window is shown in purple. **d**, Upper interface of HCN1 (Protein Data Bank (PDB) ID: 5U6O), shown in analogous view to **c**. **e**, Bound phospholipid density with its head group coordinated by R197, K200 and Y290.

C-linker snakes upwards below the S4–S5 linker, coming within 4 Å of the backbone carbonyl of the S4 helix (Fig. 3b). Mutations designed to disrupt this charge to helix-dipole interaction—R310K/Q/N/E/A—did not yield any currents, despite wild-type-like expression of R310K (Extended Data Fig. 6h, i), supporting a critical role for R310 in channel gating. The rest of the S4–S5–C-linker interaction surface appears to be formed by van der Waals contacts and potential hydrogen bonds between Y193 (in S5) and T306 (in the C-linker) as well as between R197 (in S5) and T303 (in the C-linker).

We carried out extensive mutagenesis on most of the residues that make productive interactions at the intracellular interface (Fig. 3a, b, residues coloured by effect). All mutants that generated measurable currents (K187A, D188A, R190A, F191A, T303A, R307A and R314E) require more energy for channel opening; that is, the midpoints of activation shift towards more negative potentials (Fig. 3a, Extended Data Fig. 6a), with the exception of wild-type-like N192A and R314A. Many mutants (I189A, Y193F/A, F194V/A, R197K/Q/A, K200Q/A, T306A, F309A and R310K/Q/N/E/A) did not give currents (data not shown). However, when complementary RNAs (cRNAs) encoding various loss-of-function mutations (I189A, R197K, K200Q, T306A and R310K) were individually mixed and co-injected with a gain-of-function double-mutant (Q80A/R177K) cRNA, we observed currents with activation curves shifted to the left. Such a phenotype is intermediate between those of the loss-of-function and gain-of-function mutants (Extended Data Fig. 7a, b), consistent with formation of heterotetrameric channels. This behaviour suggests that when these specific loss-of-function mutations are present in a homotetrameric 4/4 stoichiometry, they shift the channel activation potential leftwards, outside of the practical measurement range.

The mutations at the intracellular sensor–pore interface might affect the energetics of the sensor, the pore or the coupling between sensor and pore. We consider it likely that at least some of these mutants alter coupling energetics (12 mutants in total, covering the entire S4–S5–C-linker interface) (Fig. 3a, b, Extended Data Fig. 6a, b). However, owing to the technical challenge of monitoring sensor function in KAT1 mutants (by gating currents or fluorescence) we cannot conclusively determine the contributions of each individual residue to sensor–pore coupling. As a partial and preliminary readout of sensor motion and function, we used limiting-slope analyses using macroscopic currents for the two VSD residues that are at the intracellular VSD–pore interface (K187A and D188A) as a way to estimate the amount of charge moved upon channel activation. Compared with the wild-type channel, the D188A mutant moves a similar amount of charge upon activation, despite its left-shifted ionic current activation; therefore, D188A may impair sensor–pore coupling (Extended Data Fig. 7c, d). Low expression levels of K187A prevented robust limiting-slope analyses (data not shown), and it is possible that K187A impairs VSD function rather than coupling. More generally, the result that the majority of KAT1 VSD-pore mutants generate channels with an energetic bias for the closed state over the open state is consistent with the hypothesis that the pore domain of KAT1 is closed by default and the VSD performs work to open the pore at negative potentials[4,14–16]. Future functional and structural experiments conducted in the isolated pore domain of KAT1 could be used to further test this hypothesis.

Within the plane of the membrane, the VSD and pore of KAT1 are separated by a hydrophobic window. This window is absent in the HCN1 structure, in which S4 and S5 form zipper-like interactions along their length[13] (Fig. 3d). Of note, the hydrophobic window of KAT1 is filled with a tubular density (Fig. 3e, Extended Data Fig. 6e–g)—which we have putatively assigned as the alkyl chain of an intercalated phospholipid. This bound phospholipid appears in a conformation that is not observed in other ion channel structures. The head group of this intercalated phospholipid is coordinated by charge–charge and hydrogen-bonding interactions between R197 and K200 on S5 and Y290 on S6 (Fig. 3e). All mutations introduced to the lipid-coordinating residues (R197K/Q/A, K200Q/A and Y290F/A) abrogated currents (data not shown), despite wild-type-like membrane expression of R197K and K200Q mutants (Extended Data Fig. 6h, i), suggesting a structural or functional role for the bound lipid. During a molecular dynamics simulation approximately 3.5 μs in duration, in which a lipidless KAT1 was initially placed in a 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) bilayer, lipid molecules from the bulk stably occupied binding conformations similar to that seen in the cryo-EM structure (Extended Data Fig. 6e, f). KAT1 and other plant plasma membrane K_v channels are strongly modulated by PtdIns(4,5)P_2 through an unknown mechanism[17] and the lipid in the hydrophobic window may indicate a binding site of PtdIns(4,5)P_2 or some other modulatory lipid. Given the placement of this binding site at the functionally critical S4–S5–S6 interface, the bound lipid may constitute an integral component of the gating machinery. In addition, KAT1 is known to open very slowly: the time constants for gating and ionic currents are separated by approximately three orders of magnitude (time constants of gating current and ionic current activation are approximately 270 μs and 120 ms, respectively)[3]. A requirement for lipid binding (Fig. 3e) or reorientation upon gating is a speculative, but testable hypothesis to explain this kinetic disparity.

In contrast to what we observe at the intracellular interface, mutagenic perturbations at the extracellular interface formed by S1, S4 and S5 yielded mixed and nuanced effects on the channel energetics (Extended Data Fig. 6b–d). These mutations led to four distinct phenotypes: nonfunctional channels (F83A, L172A, F207A and C211A), wild-type-like channels (F81A/L, I166A and F215A), and channels that are energetically biased towards the closed state (V178A) or open state (M169A), compared with wild type. Mutations that abrogate ionic currents support the idea that the S1–S4–S5 interface might be important for channel assembly and stability. However, its role as a major pathway of energy transfer from VSD to pore remains to be defined.

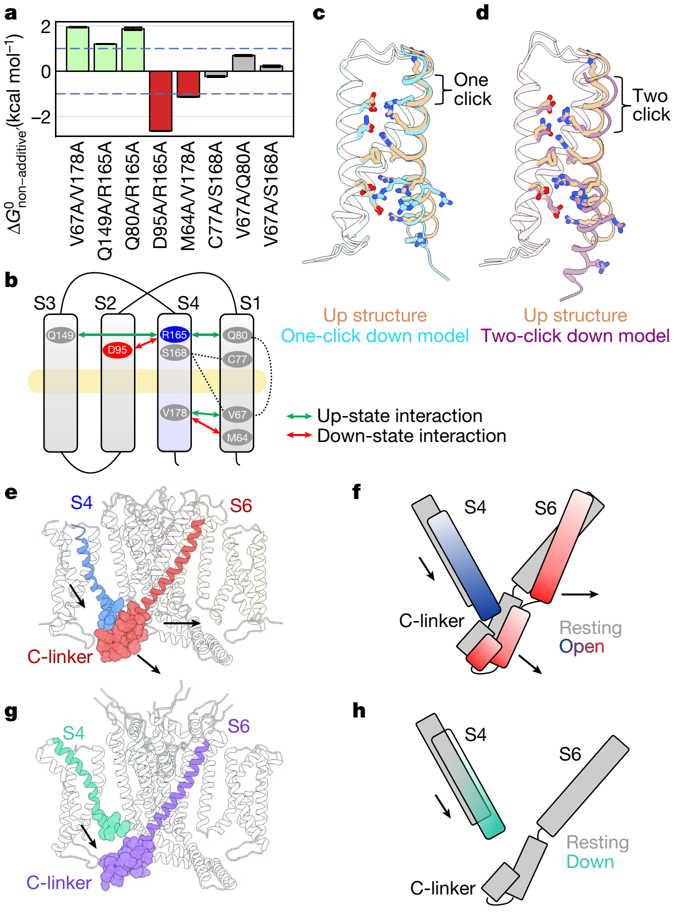


Fig. 4 | Hypothetical modelling of the KAT1 VSD in the down state, and implications for electromechanical coupling and gating polarity. **a**–**d**, Hypothetical modelling of KAT1 VSD activation. **a**, Plot of nonadditive energies derived from double-mutant cycle analysis, with 1 kcal mol^{-1} threshold shown as dotted lines. Data shown are the calculated $\Delta G_{\text{nonadditive}}$ (Methods) using the mean and s.d. of each dataset. Sample sizes are provided in Extended Data Fig. 8b (from which the data are derived). **b**, Schematic of interacting residues, using the same colour scheme as in **a**. **c**, One-click down-state model (blue) derived from interacting pairs and equilibrated by molecular dynamics simulation. **d**, Two-click down-state model (purple) derived from interacting pairs and equilibrated by molecular dynamics simulation. **e**–**h**, Hypothetical models for electromechanical coupling and gating polarity in KAT1. **e**, Side view of KAT1em, with pseudo CNBDs removed for clarity. Van der Waals sphere representation highlights tight packing between the up-state S4 and closed C-linker, packing which would be disrupted by a one-click downward movement of S4. **f**, Cartoon of S4–C-linker coupling. **g**, Side view of depolarization-activated channel rat EAG1[28], highlighting disengagement of S4 and C-linker when the sensor is up and the intracellular gate is closed. N and C-terminal cytosolic domains have been removed for clarity. **h**, cartoon of S4–C-linker coupling in rat EAG1, highlighting how the increased distance between S4 and C-linker might preclude activation by hyperpolarization.

Given the structural and energetic relationship between VSD and pore, in particular the tight packing at the S4–S5–C-linker interface and the severe loss-of-function phenotypes of mutants at this interface, we investigated how KAT1 might open upon membrane hyperpolarization. As a first step towards answering this question, we sought to estimate the extent of the conformational change in the VSD associated with KAT1 opening. We used double-mutant cycle analysis[18] (Fig. 4a–d, Extended Data Fig. 8) to investigate a subset of residue–residue interactions that might change upon hyperpolarization and construct hypothetical down-state models of the VSD, which correspond to the open channel at hyperpolarized potentials. Calculated $|\Delta G_{\text{nonadditive}}|$ values greater than 1 kcal mol^{-1} (ref. [19]) are interpreted as a state-dependent interaction between two residues, with negative values of $\Delta G_{\text{nonadditive}}$ indicating stronger interaction in the down state, and positive values indicating stronger interaction in the up state. In brief, $\Delta G_{\text{nonadditive}}$ is defined as $\Delta G_{\text{nonadditive}} = \Delta\Delta G_{\text{mut1}} + \Delta\Delta G_{\text{mut2}} - \Delta\Delta G_{\text{mut1,2}}$, in which $\Delta\Delta G_{\text{mut1}}$ is the difference in deactivation energy between single mutant 1 and wild-type, $\Delta\Delta G_{\text{mut2}}$ is the difference in deactivation energy between single mutant 2 and wild-type, and $\Delta\Delta G_{\text{mut1,2}}$ is the difference in deactivation energy between double mutant 1,2 and wild-type. On the basis of this mutant cycle analysis, we identified two residues on S4, R0 (R165) and V178, each of which exchanges different interaction partners upon VSD activation (Fig. 4a,b, Extended Data Fig. 8). Furthermore, metal bridging experiments point to a cadmium-dependent interaction between R165C (R0 on S4) and C77 (on S1) that promotes channel opening and thus represents an additional down-state interacting pair (Extended Data Fig. 9).

These down-state interacting pairs were then used to construct simplified, hypothetical 'one-click' and 'two-click' down-state VSD models, in which the S4 helix moves downward by one and two helical turns, respectively, in the context of the isolated KAT1 VSD (Fig. 4c, d). Molecular dynamics simulations were used to calculate the amount of gating charge displaced during these putative transitions, enabling comparison with electrophysiological limiting-slope estimates, which provide a lower bound of around 0.75 *e* per VSD in KAT1[3]. We obtained 1.02 *e* for the one-click model and 1.57 *e* for the two-click model by molecular dynamics simulation (Extended Data Fig. 8c). Our hypothetical models, particularly the one-click model, are consistent with the limiting-slope estimate in the literature, the double mutant cycle and metal-bridge constraints, and previous second-site suppressor studies[20,21] (Extended Data Fig. 8e). Our proposed KAT1 VSD motion would encompass a displacement of approximately 5–7 Å, similar to that proposed for depolarization-activated channels[22] and observations in other VSD structures[23–26]. Thus, the main question becomes how a canonical downward VSD motion might lead to pore opening in a hyperpolarization-activated channel[2,3,27].

In our gating model, the downward, hyperpolarization-driven movement of S4 is directly coupled to a subsequent lateral reorientation of the C-linker of the neighbouring subunit, ultimately opening the S6 gate (Fig. 4e, f). Although it shares a similar architecture to KAT1, the structure of depolarization-activated EAG1 (also captured with an up-state voltage sensor and closed intracellular gate), shows S4 to be disengaged from the C-linker[28] (Fig. 4g). According to our model, a downward movement of S4 of EAG1 would be unable to trigger channel opening upon membrane hyperpolarization (Fig. 4h), consistent with the depolarization-activated phenotype of EAG1. It is also worth noting that although KAT1 is nominally not domain-swapped, the tight interaction between S4 and the C-linker in its adjacent subunit at rest (0 mV) ultimately leads to a process of activation gating dominated by direct communication between subunits.

Our proposal for a direct coupling mechanism for KAT1 contrasts with the allosterically coupled nature of voltage-sensitive gating reported for HCN channels[5,29–31], where coupling might not be as strong as suggested for KAT1[3]. Supporting this proposal, KAT1, unlike HCN, is not activated by cyclic nucleotides[10] and the structural conformation of the KAT1em pseudo-CNBD is already compatible with an 'activated' ligand-binding domain conformation even in the absence of ligand[13] (Extended Data Fig. 4c–g). Therefore, we suggest that KAT1 is perhaps mechanistically closer to a 'reversed' depolarization-activated, non-domain-swapped channel such as EAG1 or human ERG, even though KAT1 lacks the cytoplasmic Per-Arnt-Sim domain of EAG1 and human ERG[28,32,33]. In view of these results, the present proposal is likely to have direct implications to the mechanism of gating and electromechanical coupling in non-domain-swapped channels such as EAG1 and human ERG, in which electric field transduction (and not nucleotide binding) represents the sole driving force for channel gating. We anticipate that the KAT1 structure will serve as a framework for future functional and engineering studies of ion channels. Such efforts in plants might hold promise in improving carbon assimilation and optimal biomass production[34].

Note added in proof: Since this paper was accepted, two contributions independently addressed electromechanical coupling in HCN1 channels[35,36]. These studies establish the structural and energetic underpinnings of the allosteric communication between the voltage sensors and the activation gate of HCN1. Further, they support the present conclusion that highlight the divergence between HCN channels and the direct coupling mechanism suggested here for KAT1.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2335-4.

1. Bezanilla, F. How membrane proteins sense voltage. *Nat. Rev. Mol. Cell Biol.* **9**, 323–332 (2008).
2. Männikkö, R., Elinder, F. & Larsson, H. P. Voltage-sensing mechanism is conserved among ion channels gated by opposite voltages. *Nature* **419**, 837–841 (2002).
3. Latorre, R. et al. Molecular coupling between voltage sensor and pore opening in the *Arabidopsis* inward rectifier K$^+$ channel KAT1. *J. Gen. Physiol.* **122**, 459–469 (2003).
4. Blunck, R. & Batulan, Z. Mechanism of electromechanical coupling in voltage-gated potassium channels. *Front. Pharmacol.* **3**, 166 (2012).
5. Altomare, C. et al. Integrated allosteric model of voltage gating of HCN channels. *J. Gen. Physiol.* **117**, 519–532 (2001).
6. Long, S. B., Campbell, E. B. & Mackinnon, R. Voltage sensor of Kv1.2: structural basis of electromechanical coupling. *Science* **309**, 903–908 (2005).
7. Vardanyan, V. & Pongs, O. Coupling of voltage-sensors to the channel pore: a comparative view. *Front. Pharmacol.* **3**, 145 (2012).
8. Chowdhury, S. & Chanda, B. Perspectives on: conformational coupling in ion channels: thermodynamics of electromechanical coupling in voltage-gated ion channels. *J. Gen. Physiol.* **140**, 613–623 (2012).
9. Hedrich, R. Ion channels in plants. *Physiol. Rev.* **92**, 1777–1811 (2012).
10. Hoshi, T. Regulation of voltage dependence of the KAT1 channel by intracellular factors. *J. Gen. Physiol.* **105**, 309–328 (1995).
11. Moroni, A. et al. Mutation in pore domain uncovers cation- and voltage-sensitive recovery from inactivation in KAT1 channel. *Biophys. J.* **78**, 1862–1871 (2000).
12. Hertel, B. et al. KAT1 inactivates at sub-millimolar concentrations of external potassium. *J. Exp. Bot.* **56**, 3103–3110 (2005).
13. Lee, C.-H. & MacKinnon, R. Structures of the Human HCN1 Hyperpolarization-Activated Channel. *Cell* **168**, 111–120 (2017).
14. Yifrach, O. & MacKinnon, R. Energetics of pore opening in a voltage-gated K$^+$ channel. *Cell* **111**, 231–239 (2002).
15. Soler-Llavina, G. J., Chang, T. H. & Swartz, K. J. Functional interactions at the interface between voltage-sensing and pore domains in the Shaker K$_v$ channel. *Neuron* **52**, 623–634 (2006).
16. Ledwell, J. L. & Aldrich, R. W. Mutations in the S4 region isolate the final voltage-dependent cooperative step in potassium channel activation. *J. Gen. Physiol.* **113**, 389–414 (1999).
17. Liu, K., Li, L. & Luan, S. An essential function of phosphatidylinositol phosphates in activation of plant shaker-type K$^+$ channels. *Plant J.* **42**, 433–443 (2005).
18. Carter, P. J., Winter, G., Wilkinson, A. J. & Fersht, A. R. The use of double mutants to detect structural changes in the active site of the tyrosyl-tRNA synthetase (*Bacillus stearothermophilus*). *Cell* **38**, 835–840 (1984).
19. Yarov-Yarovoy, V. et al. Structural basis for gating charge movement in the voltage sensor of a sodium channel. *Proc. Natl Acad. Sci. USA* **109**, E93–E102 (2012).
20. Lai, H. C., Grabe, M., Jan, Y. N. & Jan, L. Y. The S4 voltage sensor packs against the pore domain in the KAT1 voltage-gated potassium channel. *Neuron* **47**, 395–406 (2005).
21. Grabe, M., Lai, H. C., Jain, M., Jan, Y. N. & Jan, L. Y. Structure prediction for the down state of a potassium channel voltage sensor. *Nature* **445**, 550–553 (2007).
22. Vargas, E. et al. An emerging consensus on voltage-dependent gating from computational modeling and molecular dynamics simulations. *J. Gen. Physiol.* **140**, 587–594 (2012).
23. Li, Q. et al. Structural mechanism of voltage-dependent gating in an isolated voltage-sensing domain. *Nat. Struct. Mol. Biol.* **21**, 244–252 (2014).
24. Guo, J. et al. Structure of the voltage-gated two-pore channel TPC1 from *Arabidopsis thaliana*. *Nature* **531**, 196–201 (2016).
25. Yan, Z. et al. Structure of the Na$_v$1.4-β1 complex from electric eel. *Cell* **170**, 470–482.e11 (2017).
26. Xu, H. et al. Structural basis of Nav1.7 inhibition by a gating-modifier spider toxin. *Cell* **176**, 702–715.e14 (2019).
27. Sesti, F., Rajan, S., Gonzalez-Colaso, R., Nikolaeva, N. & Goldstein, S. A. N. Hyperpolarization moves S4 sensors inward to open MVP, a methanococcal voltage-gated potassium channel. *Nat. Neurosci.* **6**, 353–361 (2003).
28. Whicher, J. R. & MacKinnon, R. Structure of the voltage-gated K$^+$ channel Eag1 reveals an alternative voltage sensing mechanism. *Science* **353**, 664–669 (2016).
29. Chen, S., Wang, J., Zhou, L., George, M. S. & Siegelbaum, S. A. Voltage sensor movement and cAMP binding allosterically regulate an inherently voltage-independent closed-open transition in HCN channels. *J. Gen. Physiol.* **129**, 175–188 (2007).
30. Kusch, J. et al. Interdependence of receptor activation and ligand binding in HCN2 pacemaker channels. *Neuron* **67**, 75–85 (2010).
31. Alvarez-Baron, C. P., Klenchin, V. A. & Chanda, B. Minimal molecular determinants of isoform-specific differences in efficacy in the HCN channel family. *J. Gen. Physiol.* **150**, 1203–1213 (2018).
32. Wang, W. & MacKinnon, R. Cryo-EM structure of the open human ether-à-go-go-related K$^+$ channel hERG. *Cell* **169**, 422–430.e10 (2017).
33. Perissinotti, L. L. et al. Determinants of isoform-specific gating kinetics of hERG1 channel: combined experimental and simulation study. *Front. Physiol.* **9**, 207 (2018).
34. Papanatsiou, M. et al. Optogenetic manipulation of stomatal kinetics improves carbon assimilation, water use, and growth. *Science* **363**, 1456–1459 (2019).
35. Kasimova, M. A. et al. Helix breaking transition in the S4 of HCN channel is critical for hyperpolarization-dependent gating. *eLife* **8**, e53400 (2019).
36. Lee, C. H. & MacKinnon, R. Voltage sensor movements during hyperpolarization in the HCN channel. *Cell* **179**, 1582–1589 (2019).

Methods

Molecular biology and biochemistry

A DNA construct encoding amino acids M1–S502 was codon optimized for sf9 expression and synthesized by Integrated DNA Technologies. This gene was subcloned into a modified pFastBac vector containing a C-terminal 3C protease site, eGFP, and His$_8$ using restriction sites 5′ NotI and 3′ XbaI. Baculovirus was generated via the Bac-to-Bac method (Invitrogen). P0 virus was amplified once to yield P1 baculovirus, which was used to infect sf9 cells (ATCC CRL-1711) at a 1:50 v/v ratio. Cells were not tested for mycoplasma nor further authenticated. Cells were collected 36–40 h post infection, washed in phosphate-buffered saline pH 7.4, dounce homogenized in hypotonic buffer A (20 mM HEPES pH7.4, 20 mM KCl, 10 mM MgCl$_2$) and ultracentrifuged. This hypotonic lysis cycle was repeated four times and was subsequently followed by one cycle in hypertonic buffer (buffer A plus 800 mM NaCl). Membranes were resuspended in 50mM HEPES pH7.4, 200 mM KCl supplemented with 40% glycerol and flash frozen. For purification all steps were performed at 4 °C. Membranes were thawed, diluted with glycerol-free buffer and detergent-extracted in 50 mM HEPES pH7.4, 200 mM KCl, 1% *n*-dodecyl-β-D-maltopyranoside (DDM; Anatrace), 0.2% cholesteryl hemisuccinate Tris salt (CHS; Steraloids), asolectin (Sigma, crude) 0.05mg/ml for 90 min. Solubilized supernatant was isolated by ultracentrifugation and diluted with low-detergent buffer to drop DDM/CHS concentration to about 0.5% DDM, 0.1% CHS. Supernatant was batch bound to Cobalt IMAC Talon beads (clontech) for 2–3 h with 5 mM imidazole present. Beads were collected by low-speed centrifugation and washed in batch with 50 mM HEPES pH7.4, 200 mM KCl, 0.05% DDM (Anatrace), 0.01% CHS (Anatrace), 0.05 mg ml^{-1} asolectin (Avanti) and 15 mM imidazole. Beads were transferred to plastic column and further washed exchanging stepwise to buffer containing digitonin 0.05% (Millipore) and eluted in 50 mM HEPES pH7.4, 200 mM KCl, 0.05% digitonin and 250 mM imidazole. Protein was cleaved by HRV 3C protease[37] for 3 h, concentrated and subjected to size-exclusion chromatography (SEC) on a Superose 6 column (GE Healthcare) with running buffer: 50 mM HEPES pH7.4, 200 mM KCl, 0.05% digitonin, 2 mM CaCl$_2$. Peak fractions were collected and concentrated to 4–5 mg ml^{-1} (Millipore concentrator unit).

Cryo-EM analysis

Quantifoil 200-mesh 1.2/1.3 grids (Quantifoil) were plasma cleaned for 30 s in an air mixture in a Solarus Plasma Cleaner (Gatan). Grids were frozen in liquid nitrogen-cooled liquid ethane in a Vitrobot Mark IV (FEI) using the following parameters: 3.5 μl sample volume, 2.5 s blot time, blot force 3, 100% humidity, at a temperature of 22 °C and double filter papers on each side of the vitrobot.

Grids were screened on a 200 kV Talos side entry microscope (FEI) equipped with K2 summit direct detector (Gatan) using a Gatan 626 single-tilt holder. Replicate grids from the same preparation were shipped to the National Cryo-Electron Microscopy Facility at the National Cancer Institute. Grids were imaged on a Titan Krios with K2 detector (super-resolution mode) and GIF energy filter (set to 20 eV) at a nominal magnification of 130,000 corresponding to a super-resolution pixel size 0.532 Å per pixel. The dose rate was roughly 4.7 e^- pixel^{-1} s^{-1} and the exposure time was 12 s, yielding a total post-GIF dose of 38–43 e^- Å^{-2}. A total of 1,502 movies were collected using Latitude (Gatan). Data were processed using motioncor2[38], Ctffind4[39], and Relion 2[40]. A total of 1,500 particles were manually picked and classified in 2D to generate autopicking templates. Autopicking in Relion2 using a picking threshold of 0.5 gave about 120,000 particles, which were subjected to 2D classification. Some 110,000 particles were selected from good classes, and 10,000 of these particles were used to generate an initial model with C4 symmetry imposed. All 110,000 particles were then subjected to autorefinement, yielding a 4.3 Å nominal resolution map. Inspection of the two tetramers within the octamer indicates that they are nearly indistinguishable, and are related by about 45° rotation at the pCNBHD–pCNBHD interface. Classification of all 110,000 particles in C1-symmetry closely resembled the overall architecture of the C4-symmetry-imposed map, albeit with lower resolution and a slight tilt of the two micelles with respect to one another. The best two classes from the C1-symmetry job were combined, yielding about 90,000 particles, which were then subjected to autorefinement in C4-symmetry. Refinement of the octamer yielded a map that was used for model building of the cytosolic domains. Focused refinement on the tetramer and subsequently the transmembrane region of the tetramer gave a reconstruction with improved map quality supporting confident building of the transmembrane regions. Postprocessing of the focused transmembrane map was performed in Relion 2 using the star file of the K2 detector at 300 kV and a masked nominal resolution of 3.5 Å by 0.143 Fourier shell correlation (FSC) criterion was calculated[41–43]. Local resolution was calculated by ResMap[44] and particle orientation distribution calculated by Relion 2[40]. A *B*-factor of −134 was used for sharpening and visualization.

Model building

Swiss-Model[45,46] was used to generate homology models of KAT1em using human HCN1 and rat EAG1 as templates[13,28]. The human HCN1-template model was then stubbed to poly alanine using Chainsaw[47], and all loops were deleted. Secondary structural elements were rigid body fit to the density, and then refined in real space without secondary structure restraints using phenix.real_space_refine[48,49]. Subsequent manual building in Coot[50–52] registered secondary structural elements using bulky residues and built loops where appropriate. Residues that did not show side chain density were stubbed at the Cβ. Final refinement of the transmembrane and cytosolic domains were conducted independently, against the transmembrane domain-focused map or the full-molecule map, respectively. Strong non-crystallographic symmetry constraints in phenix.real_space_refine were used to immobilize the domain that was not currently being refined (that is, the cytosolic domain during the transmembrane domain-focused map refinement).

The tetramer model was generated by applying symmetry operations to the monomer in UCSF Chimera[53]. The octamer model was generated by docking two tetramers in Chimera using the fit-in-map tool. Side chains of the C helices at the octamer interface could not be assigned definite rotamers likely due to pseudo-symmetry and were stubbed at the Cβ.

Molecular biology and electrophysiology

The full-length, native KAT1 cDNA from *A. thaliana* was obtained from the *Arabidopsis* Biological Resource Center, and DNA was cloned into the pBSTA vector[54,55]. Mutations were introduced via site-directed mutagenesis and confirmed by Sanger sequencing. cRNA was synthesized using the T7 RNA expression Kit (Ambion, Invitrogen). Approximately 24 h after surgical removal from adult frogs, in accordance with animal usage protocol 71475 of the University of Chicago Institutional Animal Care and Use Committee, 50–100 ng cRNA in 50 nl RNase-free water was injected into enzymatically defolliculated oocytes. Oocytes were maintained at 18 °C in standard oocyte solution (SOS), a solution containing 10 mM HEPES pH 7.5, 100 mM NaCl, 5 mM KCl, 2 mM CaCl$_2$, 1 mM MgCl$_2$, and 50 μg ml^{-1} gentamycin.

Macroscopic currents were recorded 36–48 h post injection on a two-electrode voltage clamp setup, comprising a OC-720C (Warner Instruments), Digidata 1322A 16 bit digitizer (Axon Instruments) and a Windows XP PC running Clampex10.3. Oocytes were impaled with two 3M KCl-filled Ag/AgCl electrodes with resistances in the range 0.2–1.0 MΩ, in bath containing SOS. For each mutant, more than four recordings were obtained, each from a different oocyte. Non-expression of a mutant was determined by absence of tail currents for more than 10 oocytes, and was confirmed in an independent injection session. KAT1 K$^+$ currents were evoked by voltage steps of

1 s, going from 0 to −190 mV in 10-mV steps. The holding potential was set at 0 mV except for extremely right-shifted mutants, the holding potential was set to +20 mV or +70 mV in order to measure the full activation curve.

The isochronal tail currents were measured in isopotential condition after the decay of the oocyte linear capacitive response. The conductance–voltage relation, $G(V)$, was obtained by constrained fitting the isochronal tail current I_{TAIL} to:

$$G(V) = A_2 + \frac{(A_1 - A_2)}{1 + e^{(V - V_\mathrm{h})zF/RT}},$$

in which V_h is the half-activation voltage, R is the gas constant, T is the absolute temperature, z is the apparent gating charge, and F is Faraday's number. The first derivatives of the raw data ($I_{\mathrm{TAIL}}(V)$) curve were numerically calculated and normalized. For the majority of the mutants, a clear peak in the first derivative was observed; the mean and variance of the peak were used to constrain the calculation of V_h. In extremely left-shifted mutants in which the peak in the derivative was not experimentally observed, the last (most negative) voltage was set as the maximum value for V_h with the minima set as −300 mV. Initial values for z were set to that of the wild-type channel, and the range of possible values is 0–4. Additional information is provided in the Supplementary Methods.

Individual $G(V)$ relations were fitted using maximum likelihood by the Monte-Carlo Markov Chain method in the lmfit package (https://lmfit.github.io/) in Python. The G/G_{MAX} curve was obtained from normalizing the $G(V)$ by A1 and A2 values from the fit. A Bayesian sampling of the posterior distribution for the parameters V_h and z applied to the normalized dataset shows single solutions for all the mutants. Recordings were excluded from analysis if leak or endogenous currents prevented analysis. A record was determined to be an outlier and thus excluded, if the V_h was more than 10 mV (approximately two standard deviations) outside the mean of the normalized ensemble, or if z was more than two standard deviations outside the mean of the normalized ensemble. In all figures data are presented as mean values, with a surrounding area depicting standard deviation.

For metal bridging experiments, oocytes were recorded in SOS solution supplemented with 100 µM EDTA. After taking an initial recording in SOS + 100 µM EDTA, the solution was exchanged to SOS + 100 µM $CdCl_2$, another recorded taken, and the solution again exchanged to SOS + 100 µM EDTA and a final record taken, the whole process performed on the same oocyte. This process was then biologically replicated five times (five different oocytes), and representative currents from one oocyte are shown.

Double-mutant cycle analysis

Three types of residue-residue pair were selected by visual inspection of the structure: up-state pairs, down-state pairs, and negative control pairs (residues with interactions that are expected to be similar in both states). Data were processed as in the section above, and $\Delta G_{\mathrm{o}\to\mathrm{c}}$ (deactivation energy) values extracted. These $\Delta G_{\mathrm{o}\to\mathrm{c}}$ values were then used to calculate $\Delta G_{\mathrm{nonadditive}}$ as follows:

$$\Delta G_{o\to c} = -zFV_\mathrm{h}$$
$$\Delta\Delta G_{\mathrm{mut}} = \Delta G_{o\to c}^{\mathrm{wt}} - \Delta G_{o\to c}^{\mathrm{mut}} = -z^{\mathrm{wt}}FV_\mathrm{h}^{\mathrm{wt}} + z^{\mathrm{mut}}FV_\mathrm{h}^{\mathrm{mut}}$$
$$\Delta G_{\mathrm{nonadditive}} = \Delta\Delta G_{\mathrm{mut1}} + \Delta\Delta G_{\mathrm{mut2}} - \Delta\Delta G_{\mathrm{mut1,2}}$$

Terms are defined as follows: $\Delta G_{\mathrm{o}\to\mathrm{c}}$ is the deactivation energy. $\Delta\Delta G_{\mathrm{mut1}}$ is the difference in deactivation energy between single mutant 1 and wild type, $\Delta\Delta G_{\mathrm{mut2}}$ is the difference in deactivation energy between single mutant 2 and wild type, and $\Delta\Delta G_{\mathrm{mut1,2}}$ is the difference in deactivation energy between double mutant 1,2 and wild type. $\Delta G_{\mathrm{nonadditive}}$ then represents the extent of nonadditivity between the effect of single mutants 1 and 2 individually versus in the context of the double mutant 1,2.

Residue–residue pairs for which the magnitude of $\Delta G_{\mathrm{nonadditive}}$ was greater than 1 kcal mol^{-1} were considered to interact, and were used in modelling. The selection of the 1 kcal mol^{-1} threshold is based on previous double-mutant-cycle work[19].

Limiting-slope analysis of KAT1 channels

The ionic currents were recorded using the cut-open oocyte technique. The extracellular solution contained (in mM) 120 K-MES, 2 $CaCl_2$, 10 HEPES, pH 7.4. The intracellular solution contained (in mM) 120 K-MES, 2 EGTA, 10 HEPES, pH 7.4. The slow hyperpolarization was elicited with a voltage ramp from 0 to −100mV (1 mV s^{-1}). The inward current was fitted using cubic spline interpolation, and linear leakage correction was performed offline using a piecewise linear fitting from the beginning of the curve to the first turning point, obtained from the second derivative of the curve. Conductance-voltage relations combinations by dividing the current by the driving force, and the limiting slope (z) obtained by linear regression to the logarithm of $G(V)$ curve constrained by first and second turning points from the current second derivative. Additional information is provided in the Supplementary Methods.

Oocyte membrane expression test and confocal imaging

Oocytes for each construct (wildtype, R197K, K200Q, Y290F, and R310K) were injected as described above. After 48 h, wild-type oocytes were recorded and confirmed to give 1–2 µA of tail current. Then, 10 oocytes for each construct, as well as 10 uninjected oocytes, were washed in SOS, mechanically lysed in hypotonic lysis buffer A via pipette tip aspiration. Lysate was cleared of debris by centrifugation (10 min, 1,000*g*), and the supernatant was isolated and ultracentrifuged (30 min, 100,000*g*). The resulting membrane pellet was resuspended in 40 µl extraction buffer (50 mM HEPES pH 7.4, 200 mM KCl, 1.5% DDM, 0.3% CHS), rotated at 4 °C for 90 min and subsequently cleared by centrifugation (30 min, 12,000*g*). Supernatant was then subjected to SDS–PAGE followed by in-gel GFP imaging using a ChemiDoc Imaging System (BioRad).

For confocal imaging, oocytes were first injected and expression confirmed by recording a subset as above. Oocytes submerged in SOS were placed in a glass bottom dish (MatTek), and imaged in an Olympus DSU spinning disk confocal microscope using a 10× objective. Regions of the animal (dark) pole were imaged to avoid intrinsic autofluorescence of the vegetal (light) pole. Each sample received identical GFP channel exposures (5 s) and DIC exposures (47 ms). Images were batch normalized in SlideBook6 (3i) to allow for a fair comparison between samples, and GFP images were false-coloured in ImageJ[56].

System construction and molecular dynamics simulations

The deposited tetramer model was prepared for molecular dynamics simulations by using Coot to manually build the missing S3–S4 loop, and selecting rotamers for stubbed residues to avoid clashes. This model was then embedded into a POPC lipid bilayer solvated with a salt solution of 100 mM KCl. The symmetry axis of the protein was aligned along the z-axis. Three K^+ ions were placed at the selectivity filter ion binding sites: S0, S2 and S4 of the selectivity filter, separated by two additional water molecules occupying the binding sites S1 and S3. The final system was in an electrically neutral state with orthorhombic periodic box dimensions of about 126 × 126 × 142 $\mathrm{Å}^3$, consisting of about 227,000 atoms.

First, the all-atom system of the full channel was energy minimized for 5,000 steps, followed by a 100-ns equilibration simulation with gradually decreasing harmonic restraints being applied to the protein and the K^+ ions and the oxygen atoms of water in the selectivity filter. Then, a further 400-ns simulation was carried out with all restraints being removed. After this, the equilibrated system was simulated longer, up to 3 µs, to study the spontaneous binding of lipids to the VSD-pore interface using the special-purpose supercomputer ANTON2[57].

An isolated VSD (residues 50–189) was used to estimate the gating charge, ΔQ, corresponding to the conformational change of the VSD between different states by calculating the average displacement charge, $\langle Q_{\mathrm{d}} \rangle$, of each system. The one-click down and two-click down homology models of the KAT1 VSD were built using the program MODELLER[58], by shifting the S4 helix 3 and 6 residues downwards, respectively, from the up-state VSD in the cryo-EM structure, according to the click model of VSD movement proposed from the structural study of Ci-VSD[23], which was consistent with the classic helical-screw or sliding helix model.

The up-state VSD was inserted into a pure POPC lipid bilayer and the z-coordinates of the Cα atoms of the two aromatic residues F111 and F155 were used to adjust the position of the VSD along the normal axis of the membrane, which was then solvated in a 100 mM KCl solution. The final neutralized system contained about 31,000 atoms. The one-click down and two-click down systems were constructed by only replacing the up-state VSD protein with the one-click down and the two-click down VSD proteins, respectively. Thus, the three VSD systems had exactly the same size and components, with different protein conformations.

Each VSD system was energy minimized for 5,000 steps and equilibrated for 20 ns with the restraints applied on the protein been gradually decreased from 5 to 0 kcal mol Å^{-2} at 0 mV. The equilibrated systems were then simulated at −300 mV, −150 mV, 0 mV, 150 mV, and 300 mV for 50 ns. Snapshots from the last 40-ns trajectories were used to calculate the average displacement charge of each system at different transmembrane voltages, using the partial charge and unwrapped z coordinate of all the atoms[59]. The offset constant between the linearly fitted $\langle Q_{\mathrm{d}} \rangle$ of the systems was the gating charge associated with the conformational change between different states.

All the systems were built using the program VMD[60], and all the MD simulations other than the ANTON2 simulation were performed with the program NAMD[61]. The CHARMM36 force field[62,63] was used for proteins, phospholipids and ions, and the TIP3P model[64] for water molecules in both NAMD and ANTON2 simulations. All simulations were carried out in an NPT ensemble (300 K, 1 atm) with periodic boundary conditions and a time step of 2 fs. In the NAMD simulations, the temperature and pressure were constrained using the Langevin dynamics and the Nose–Hoover Langevin piston method[65,66], respectively. The electrostatic force was calculated with the particle-mesh Ewald method[67], and the van der Waals interaction was smoothly switched off at 10–12 Å. An electric field scaled by cell basis vectors was applied along the z-axis to simulate the membrane potential[68]. In the ANTON2 simulation, the temperature and pressure were constrained using the Nose–Hoover thermostat and the semi-isotropic MTK barostat[65,69]. Long-range electrostatic interactions were calculated using the k-space Gaussian split Ewald method[70].

Figure preparation

Structural figures were prepared with ChimeraX[71] and Chimera[53], with the aid of Segger[72,73], and MOLE[74].

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Cryo-EM density maps of KAT1 have been deposited in the Electron Microscopy Data Bank under accession codes EMD-21019 (full molecule) and EMD-21018 (transmembrane-focused refinement). The atomic models of the KAT1 tetramer and octamer have been deposited in the Protein Data Bank under accession code 6V1X and 6V1Y, respectively. All other data are available upon reasonable request to the corresponding author.

37. Shaya, D. et al. Voltage-gated sodium channel (NaV) protein dissection creates a set of functional pore-only proteins. *Proc. Natl Acad. Sci. USA* **108**, 12313–12318 (2011).
38. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
39. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
40. Kimanius, D., Forsberg, B. O., Scheres, S. H. W. & Lindahl, E. Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *eLife* **5**, 1–21 (2016).
41. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).
42. Chen, S. et al. High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy* **135**, 24–35 (2013).
43. Scheres, S. H. W. & Chen, S. Prevention of overfitting in cryo-EM structure determination. *Nat. Methods* **9**, 853–854 (2012).
44. Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11**, 63–65 (2014).
45. Biasini, M. et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **42**, W252-8 (2014).
46. Arnold, K., Bordoli, L., Kopp, J. & Schwede, T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**, 195–201 (2006).
47. Stein, N. CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. *J. Appl. Cryst.* **41**, 641–643 (2008).
48. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
49. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D* **74**, 531–544 (2018).
50. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
51. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
52. Brown, A. et al. Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta Crystallogr. D* **71**, 136–153 (2015).
53. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
54. Shin, T. M., Smith, R. D., Toro, L. & Goldin, A. L. High-level expression and detection of ion channels in *Xenopus* oocytes. *Methods Enzymol.* **293**, 529–556 (1998).
55. Carvalho-de-Souza, J. L. & Bezanilla, F. Nonsensing residues in S3-S4 linker's C terminus affect the voltage sensor set point in K^+ channels. *J. Gen. Physiol.* **150**, 307–321 (2018).
56. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
57. Shaw, D. E. et al. Anton, a special-purpose machine for molecular dynamics simulation. *Commun. ACM* **51**, 91–97 (2008).
58. Šali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815 (1993).
59. Khalili-Araghi, F. et al. Calculation of the gating charge for the K_v1.2 voltage-activated potassium channel. *Biophys. J.* **98**, 2189–2198 (2010).
60. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38, 27–28 (1996).
61. Phillips, J. C. et al. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
62. Best, R. B. et al. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain χ(1) and χ(2) dihedral angles. *J. Chem. Theory Comput.* **8**, 3257–3273 (2012).
63. Klauda, J. B. et al. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B* **114**, 7830–7843 (2010).
64. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
65. Martyna, G. J., Tobias, D. J. & Klein, M. L. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* **101**, 4177–4189 (1994).
66. Feller, S. E., Zhang, Y., Pastor, R. W. & Brooks, B. R. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* **103**, 4613–4621 (1995).
67. Essmann, U. et al. A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593 (1995).
68. Roux, B. The membrane potential and its representation by a constant electric field in computer simulations. *Biophys. J.* **95**, 4205–4216 (2008).
69. Martyna, G. J., Klein, M. L. & Tuckerman, M. Nosé–Hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys.* **97**, 2635–2643 (1992).
70. Shan, Y., Klepeis, J. L., Eastwood, M. P., Dror, R. O. & Shaw, D. E. Gaussian split Ewald: a fast Ewald mesh method for molecular simulation. *J. Chem. Phys.* **122**, 54101 (2005).
71. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).
72. Pintilie, G. D., Zhang, J., Goddard, T. D., Chiu, W. & Gossard, D. C. Quantitative analysis of cryo-EM density map segmentation by watershed and scale-space filtering, and fitting of structures by alignment to regions. *J. Struct. Biol.* **170**, 427–438 (2010).
73. Pintilie, G., Chen, D.-H., Haase-Pettingell, C. A., King, J. A. & Chiu, W. Resolution and probabilistic models of components in CryoEM maps of mature P22 bacteriophage. *Biophys. J.* **110**, 827–839 (2016).
74. Pravda, L. et al. MOLEonline: a web-based tool for analyzing channels, tunnels and pores (2018 update). *Nucleic Acids Res.* **46** (W1), W368–W373 (2018).
75. James, Z. M. & Zagotta, W. N. Structural insights into the mechanisms of CNBD channel function. *J. Gen. Physiol.* **150**, 225–244 (2018).
76. Afonine, P. V. et al. New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr. D* **74**, 814–840 (2018).

77. Smart, O. S., Neduvelil, J. G., Wang, X., Wallace, B. A. & Sansom, M. S. P. HOLE: a program for the analysis of the pore dimensions of ion channel structural models. *J. Mol. Graph*. **14**, 354–360 (1996).

Acknowledgements We thank U. Baxa and T. J. Edwards at NCEF for cryo-EM data collection; P. Rodriguez, J. Austin II and T. Lavoie at the University of Chicago Advanced Electron Microscopy Facility for microscope maintenance and training; and P. Bezanilla, M. Zhao, T. Li and B. Reddy for advice and discussions at all stages of the project. Anton 2 computer time was provided by the Pittsburgh Supercomputing Center (PSC) through Grant R01GM116961 from the National Institutes of Health. The Anton 2 machine at PSC was generously made available by D. E. Shaw Research. This research was partly supported by the National Cancer Institute's National Cryo-EM Facility at the Frederick National Laboratory for Cancer Research. This work was supported by the Consortium for Membrane Protein Dynamics and by grant R01GM088406 to EP. MDC was supported by F30MH116647 and T32GM007281.

Author contributions M.D.C. and E.P. conceived the project. M.D.C. performed structural, biochemical and most electrophysiological experiments. G.F.C. performed limiting-slope analyses and processed all electrophysiology data. R.S. constructed down-state models and ran simulations. All authors contributed to manuscript preparation.

Competing interests The authors declare no competing interests.

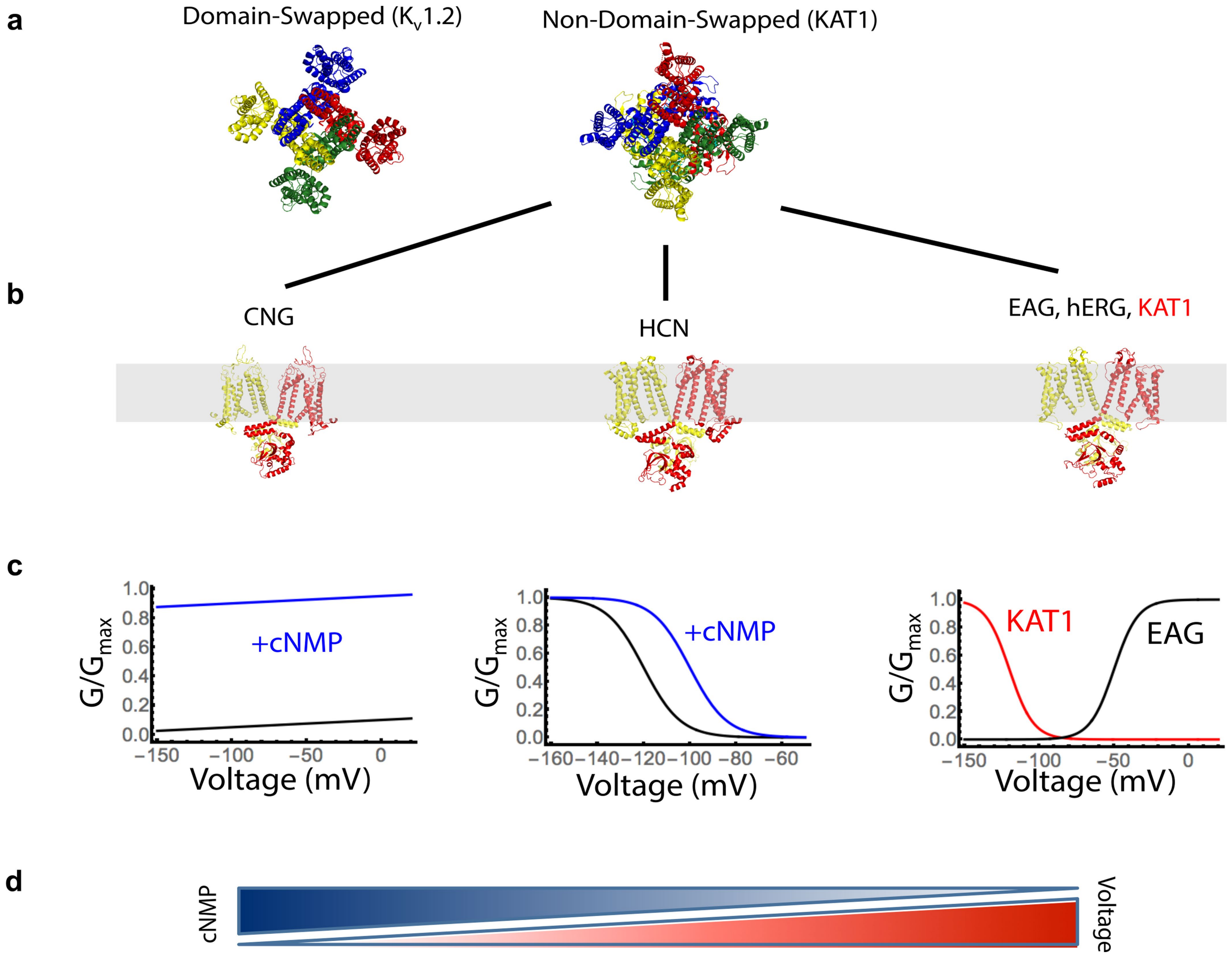
Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2335-4.

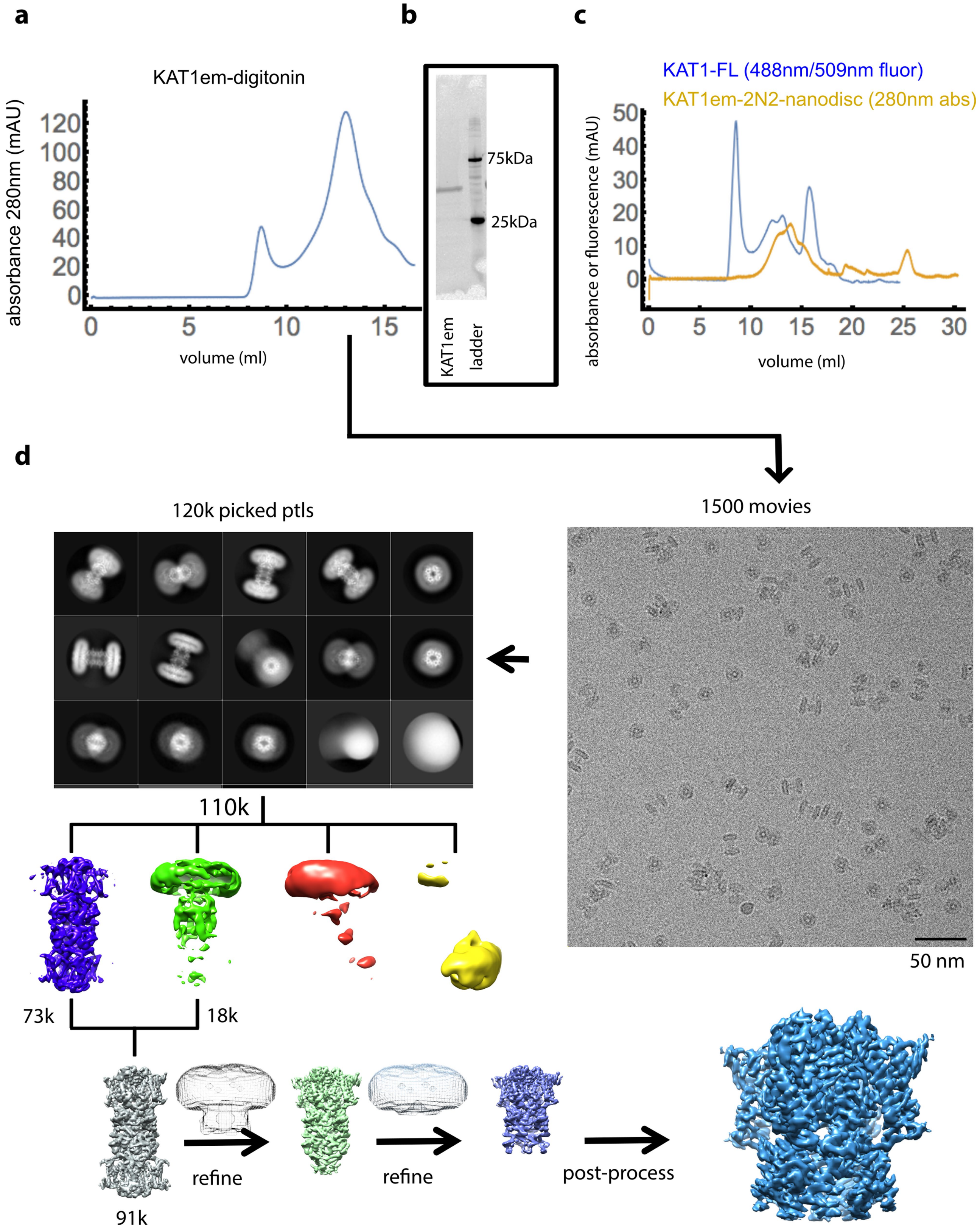
Correspondence and requests for materials should be addressed to E.P.

Peer review information *Nature* thanks Sergei Noskov, Ingeborg Schmidt-Krey and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

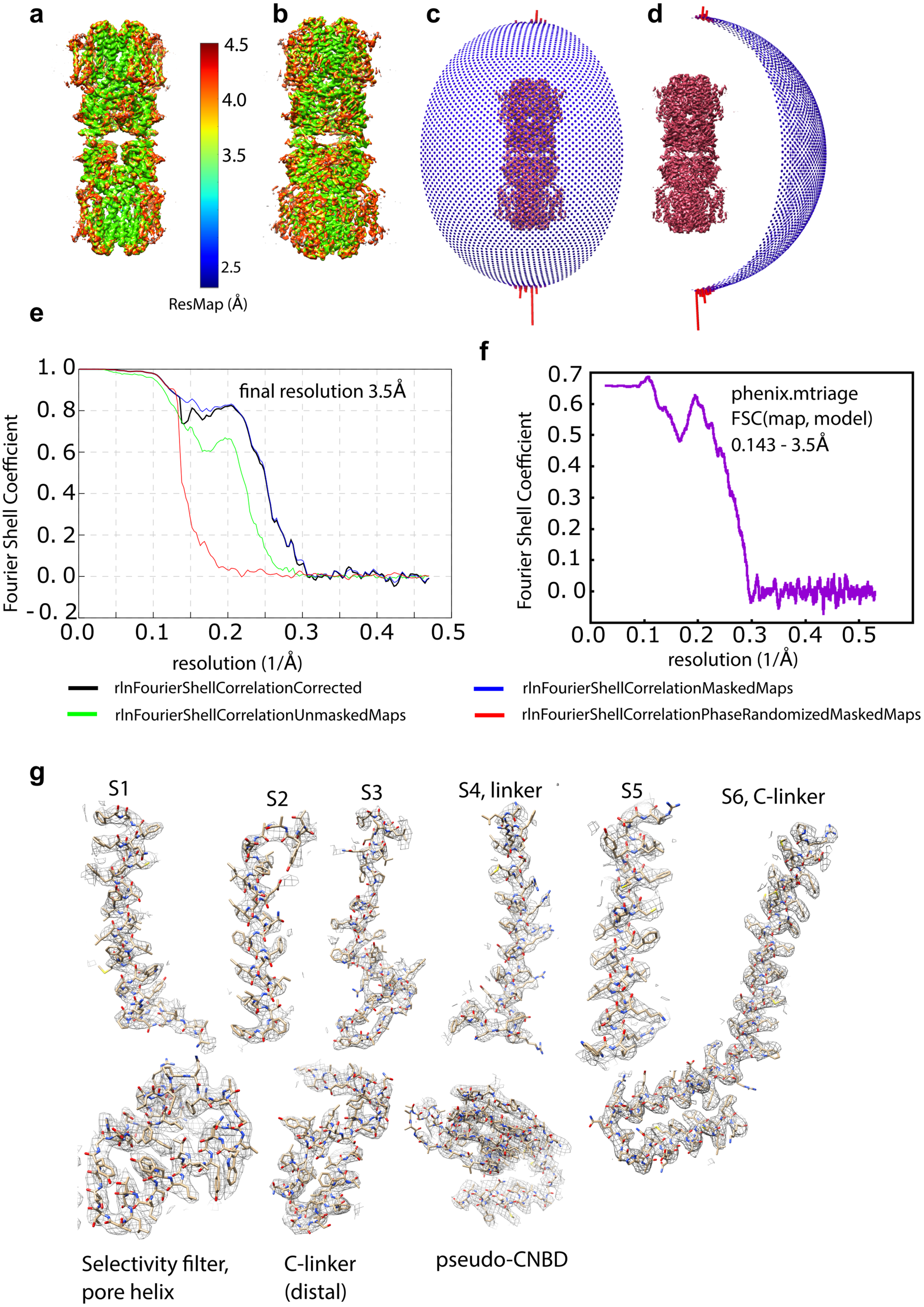
Reprints and permissions information is available at http://www.nature.com/reprints.



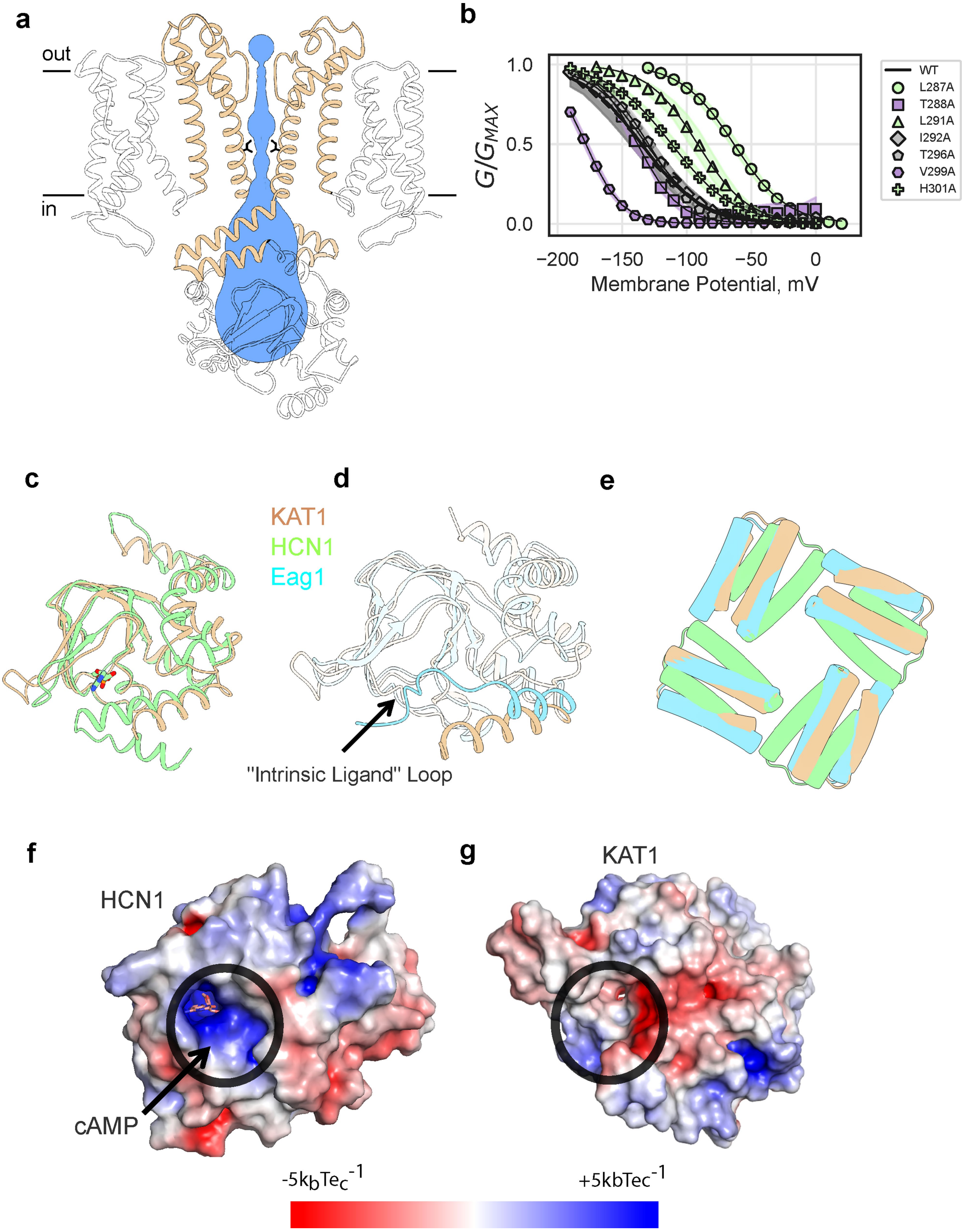
Extended Data Fig. 1 | Structural and functional diversity of tetrameric ion channels. a, Two major classes of channels, domain-swapped and non-domain-swapped are distinguished by the relative positions of voltage-sensing and pore domains. **b**, Solved structures of non-domain-swapped ion channels, two subunits shown for clarity **c**, *G*(*V*) relations of each channel subclass **d**, Gradient depiction of cyclic-nucleotide and voltage sensitivity for subclass members. Figure inspired by ref. [75].



Extended Data Fig. 2 | KAT1em biochemistry and cryo-EM workflow. a, SEC of KAT1em purified in digitonin, run on a Superose 6 column. **b**, Stain-free SDS–PAGE of purified KAT1em. SEC and SDS–PAGE results correspond to the preparation used for imaging (**d**) and are representative of three independent purifications. **c**, SEC of KAT1em in 2N2 nanodiscs (yellow trace), showing putative octamer, tetramer and empty nanodisc. Fluorescence detection SEC of full-length KAT1–GFP (blue trace) showing putative octamer and tetramer. These two samples were not subjected to any cryo-EM experiments, and are included only for the purpose of comparison. **d**, KAT1em cryo-EM workflow. From 1,500 movies, 120,000 particles were picked and subjected to 2D classification, which then yielded 110,000 particles, which were classified in 3D without imposing symmetry (4 coloured classes). Particles from the best two classes (blue and green classes, 91,000 total) were subsequently refined, imposing C4 symmetry, and using successive masks to focus on one of the tetramers and finally on the transmembrane region of one of the tetramers. Additional details are given in Methods.

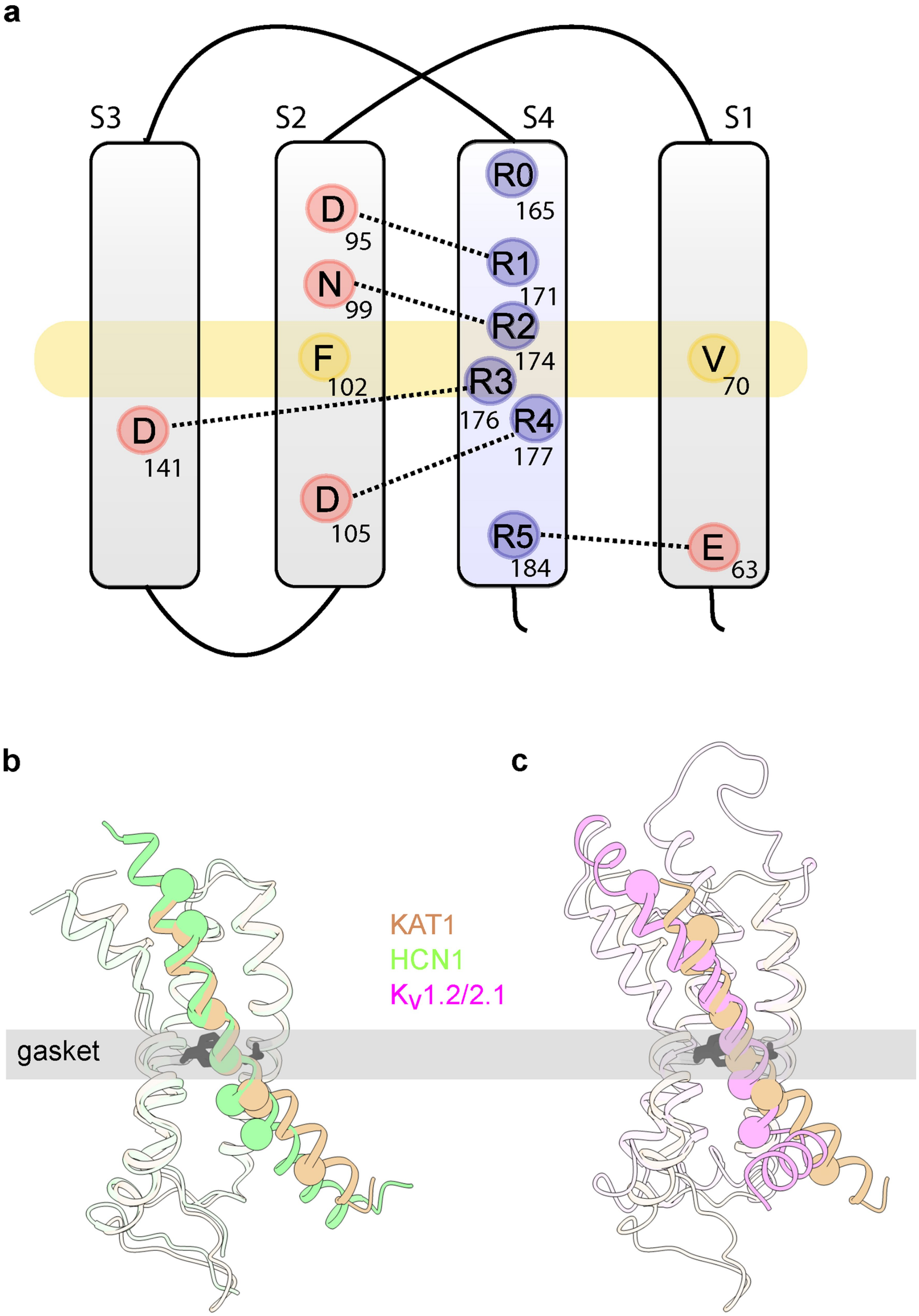


Extended Data Fig. 3 | Cryo-EM map and model validation. **a**, ResMap colouring of unfiltered half map of full molecule **b**, Same ResMap colouring as **a** on sharpened full molecule map. **c**, **d**, Ninety-degrees-rotated angular-distribution plots for refined full molecule. **e**, FSC plot for map focused on the transmembrane region. FSC 0.143 criterion is used for resolution determination[41]. **f**, FSC (map and model) plot from phenix.mtriage[76], indicating correspondence of tetramer atomic model to transmembrane domain-focused-refined density map. **g**, Details of sharpened cryo-EM density map are shown with fitted atomic model.

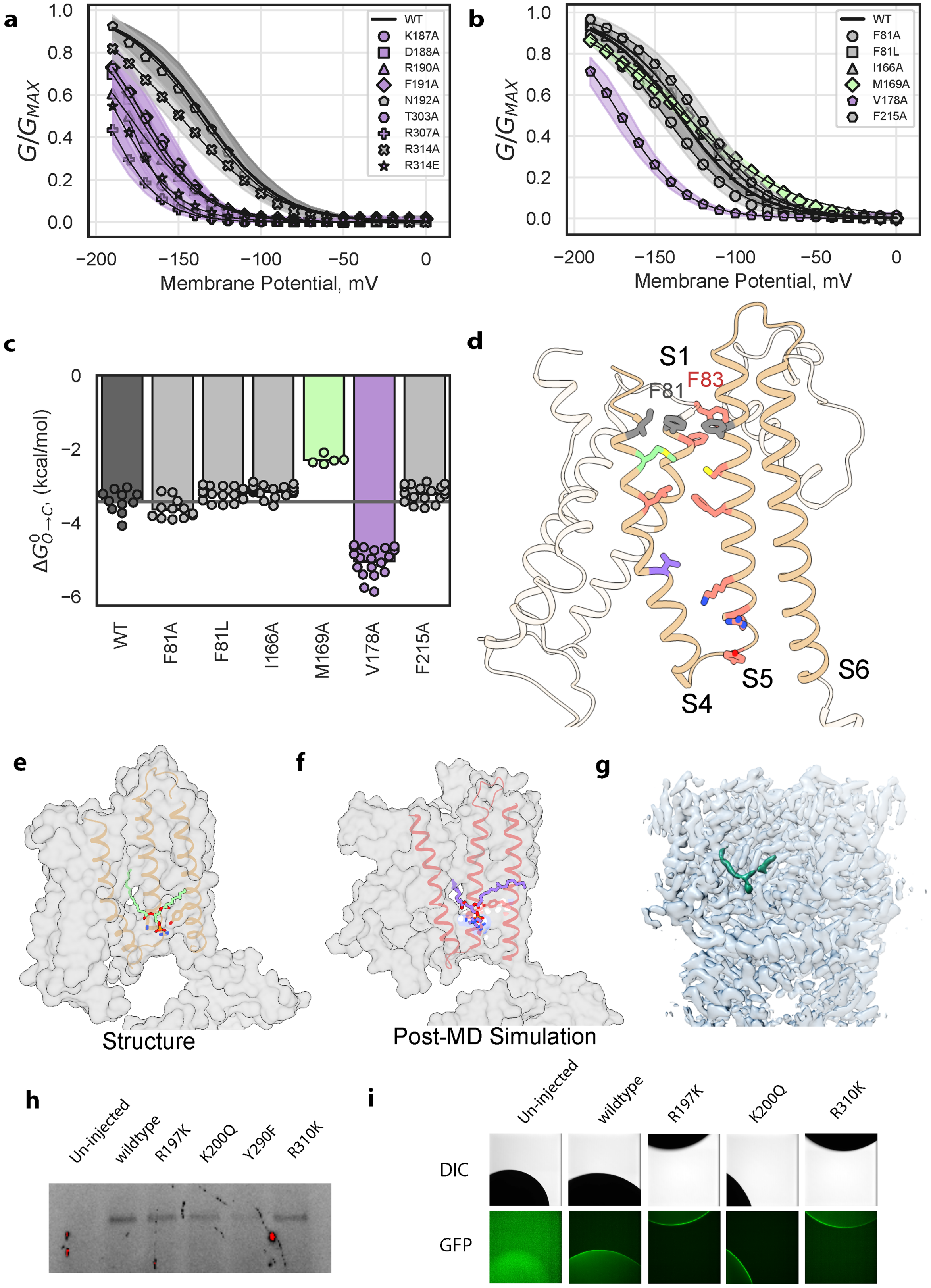


Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | KAT1em pore domain and pseudo cyclic nucleotide-binding domain. a, Side view of pore, with only two subunits shown for clarity. Permeation pathway is shown in blue, with inner gate radius calculated by MOLE[74] (1.4 Å) or HOLE[77] (1 Å), inner gate-forming I292 side chains shown as sticks. **b**, *G*(*V*) relations of pore alanine scan. Shaded error regions represent s.d., surrounding the symbols which represent the mean. Wild type (*n* = 11), L287A (*n* = 19), T288A (*n* = 4), L291A (*n* = 10), I292A (*n* = 10), T296A (*n* = 10), V299A (*n* = 8) and H301A (*n* = 10) are shown; *n* is the number of biologically independent cells. **c**, Overlay of KAT1em pseudo-CNBD (tan) and holoHCN1 CNBD (green, PDB ID: 5U6P). The ligand, HCN1-cAMP is shown as sticks in the cAMP-binding pocket. **d**, Overlay of KAT1em (tan) and EAG1 (blue, PDB ID: 5K7L). KAT1 lacks the 'intrinsic ligand' loop of EAG1. **e**, Top-down view of KAT1em (tan), holo HCN1 (green) and EAG1 (blue) overlay. Structures were aligned and superimposed on the basis of the transmembrane helices. Only C-linker hairpins are shown for clarity to compare relative rotation of the C-linker to the transmembrane domain for each structure. The relative rotation of the KAT1 C-linker matches that of EAG1 and not HCN1. **f**, **g**, Surface electrostatic potential of HCN1 (**f**) and KAT1 (**g**), respectively. Ligand-binding pockets are circled in black. KAT1 lacks a deep electropositive (blue) pocket as seen in HCN1.

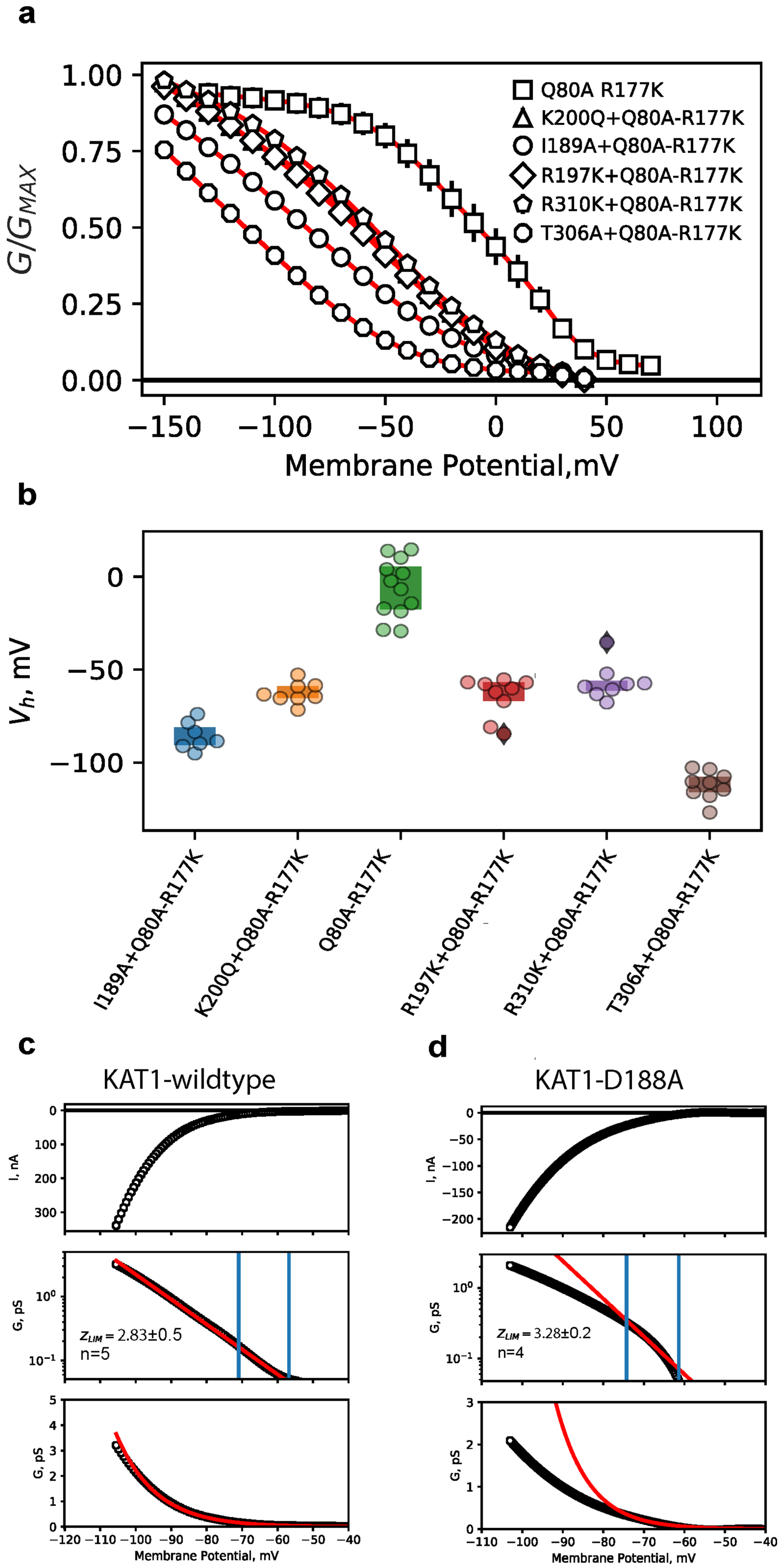


Extended Data Fig. 5 | The voltage-sensing domain of KAT1em in the up conformation. a, Diagram of key VSD features, showing hydrophobic gasket (F102 and V70, yellow) as well as all S4 charges (blue) and distributed countercharges (or counter dipoles) (red). Dashed lines indicate likely interactions in the up conformation. **b**, **c**, Overlays of KAT1em (tan) with HCN (green, PDB ID: 5U6O) and K_v1.2/2.1 (pink, PDB ID: 2R9R), respectively, highlighting structural differences between S4 helices. Cα atoms of the positively charged residues of S4 are shown as spheres.



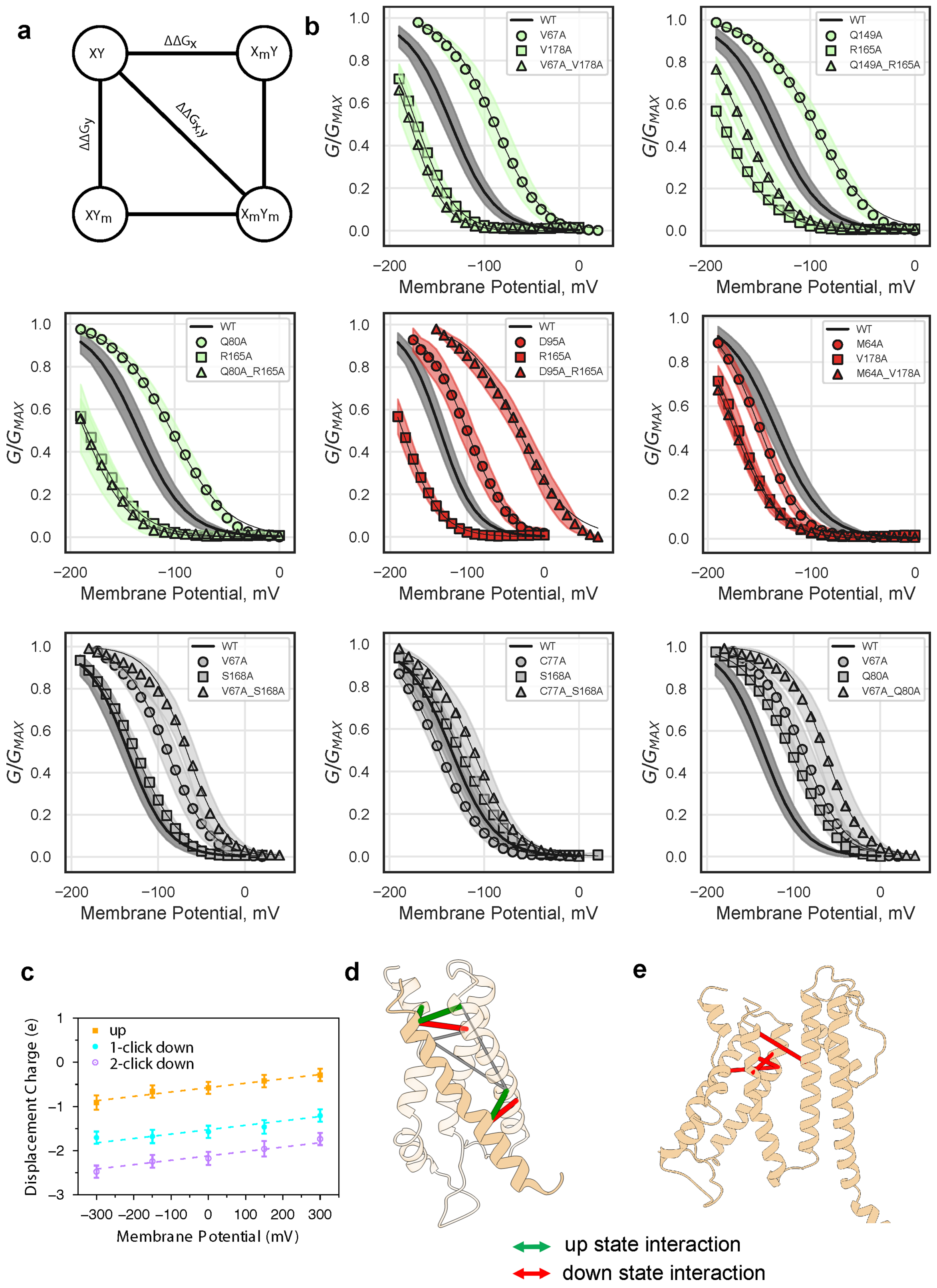
Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Structural and functional characterization of KAT1 VSD–pore interfaces. **a**, *G*(*V*) relations of S4–S5–C-linker interfacial mutants. Wild type ($n = 11$) and K187A ($n = 8$), D188A ($n = 9$), R190A ($n = 6$), F191A ($n = 12$), N192A (9), T303A ($n = 13$), R307A ($n = 14$), R314A ($n = 31$) and R314E ($n = 9$) mutants are shown; *n* is the number of biologically independent cells. Shaded regions represent s.d. and symbols represent the mean. **b**, *G*(*V*) relations of upper-interface mutants. Wild type ($n = 11$) and F81A ($n = 12$), F81L ($n = 19$), I166A ($n = 18$), M169A ($n = 5$), V178A ($n = 19$) and F215A ($n = 19$) mutants are shown; *n* is the number of biologically independent cells. Shaded regions represent s.d. and symbols represent the mean. **c**, Deactivation energies of upper-interface mutants calculated from *G*(*V*) relations in **b** (same sample sizes). **d**, Mapping of upper-interface functional data (shown in **c**). Displayed as sticks are key residues on S1: F80, F81, F83, key S4 residues: I166, M169, L172, V178, and key S5 residues: Y193, R197, K200, F207, C211, F215. **e**, **f**, Comparison of similar lipid-binding conformations observed in the structure (**e**) and after about 3.5 µs molecular dynamics simulation (**f**). **g**, Cryo-EM density map, with one bound lipid coloured green, contoured at the same contour level as the full map. **h**, SDS–PAGE and GFP in-gel imaging of *Xenopus* oocyte membrane fractions, extracted in gentle detergent (Methods). The experiment was performed once and each lane is derived from ten cells. **i**, Confocal imaging of *Xenopus* oocyte animal poles expressing various GFP-tagged constructs. Imaging was performed in a single session with normalized exposure times, and each image is representative of five independent oocytes.



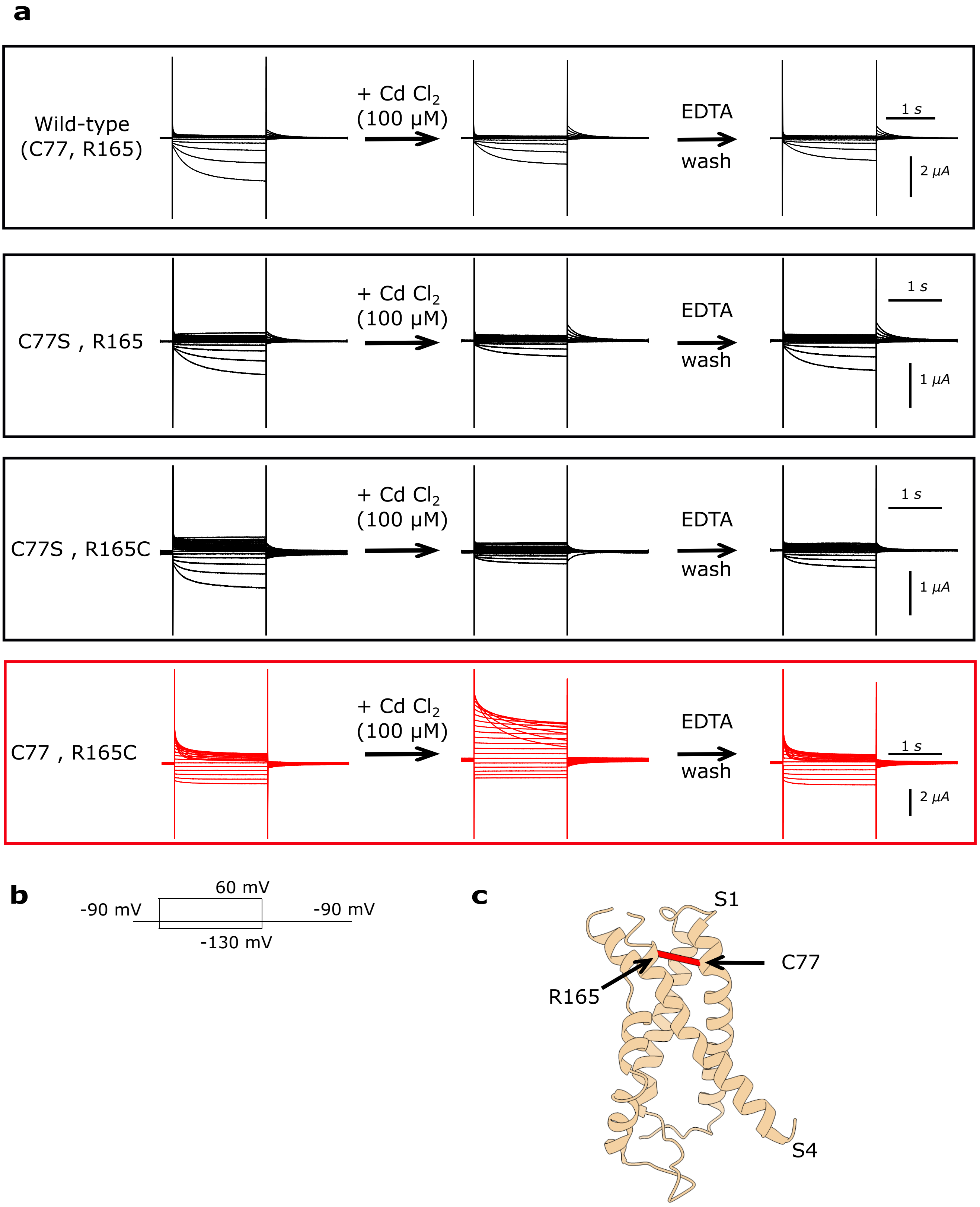
Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Detailed functional characterization of selected VSD–pore interface mutants. a, *G*(*V*) relations for cRNA mixing-coinjection experiments. cRNA encoding loss-of-function mutants (I189A, R197K, K200Q, T306A and R310K), for which no currents were observed were selected. These loss-of-function mutant cRNAs were each individually mixed with cRNA encoding a gain-of-function double mutant (Q80A–177K). Data are mean ± s.e.m. Q80A–R177K ($n = 12$), I189A + Q80A–R177K ($n = 7$), K200Q + Q80A–R177K ($n = 8$), R197K + Q80A–R177K ($n = 9$), R310K + Q80A–R177K ($n = 8$) and T306A + Q80A–R177K ($n = 9$) are shown. **b**, Plot of activation midpoints (V_h) of *G*(*V*) relations shown in **a**. **c**, **d**, Limiting-slope analyses for wild-type KAT1 (**c**) and D188A (**d**). Top, raw currents evoked by voltage ramp protocol. Middle, conductance–voltage relations, with conductance plotted on a log scale. Data points are black, fits are red. Blue vertical lines mark the first and second inflection points of the curve, the region between which was used to calculate limiting-slope (z) values (Methods). Wild type, $z = 2.83 \pm 0.5$; D188A, $z = 3.28 \pm 0.2$. data are mean ± s.d. Bottom, data (black) and fits (red) on a linear scale. n is the number of biologically independent cells.



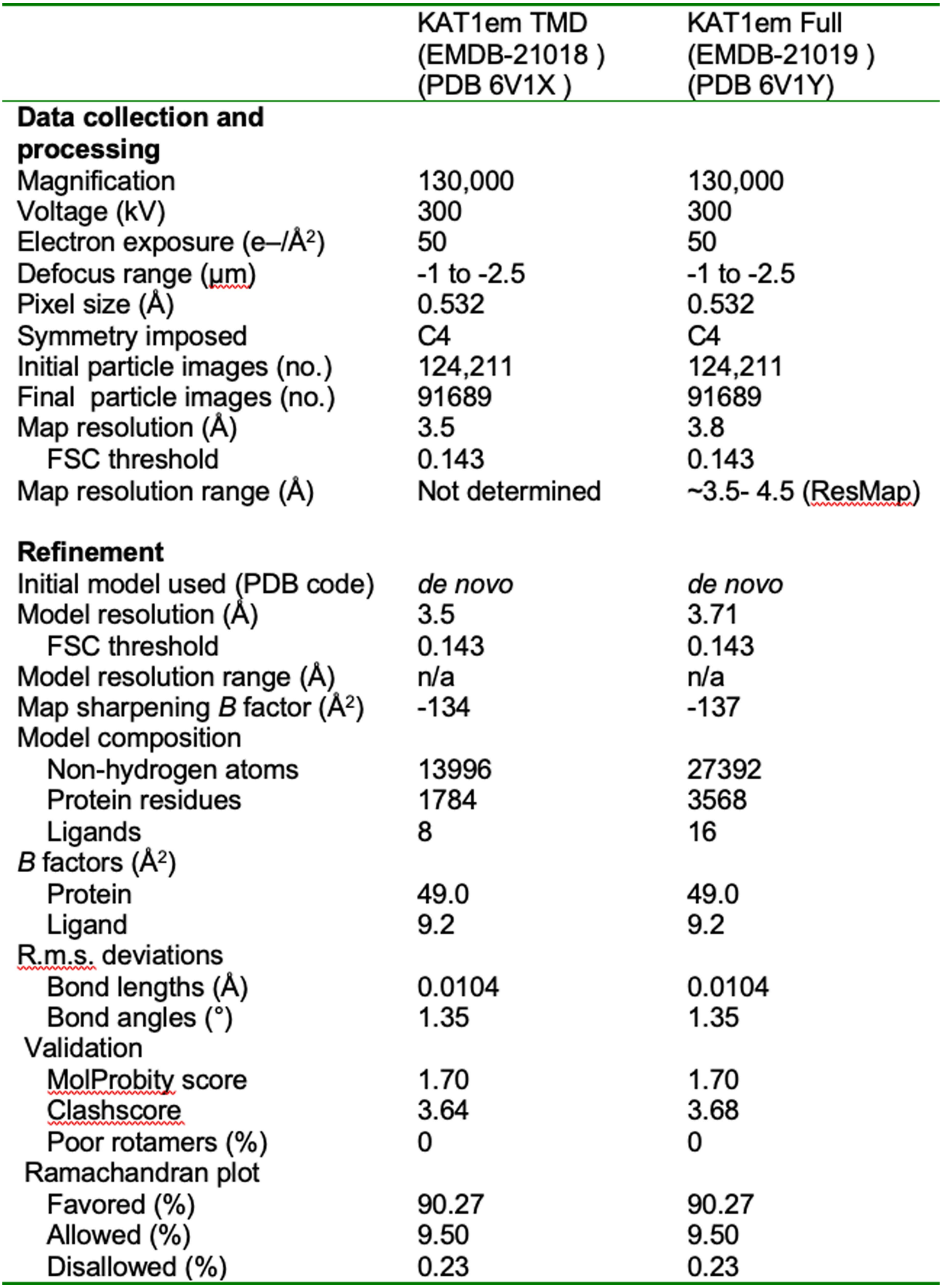
Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | VSD movement during gating. **a**, Schematic of double-mutant cycle analysis. The difference between $\Delta\Delta G_{x,y}$ and the quantity ($\Delta\Delta G_x + \Delta\Delta G_y$) determines the extent of differential interaction between residues *x* and *y* in the up and down states. **b**, *G*(*V*) relations for single and double mutants, illustrating residue–residue pairs displaying additivity (gray) and non-additivity in different directions (green, up-state interaction; red, down-state interaction). Shaded regions represent s.d. and symbols represent the mean. Wild type ($n = 11$) and M64A ($n = 11$), V67A ($n = 33$), C77A ($n = 15$), Q80A ($n = 11$), D95A ($n = 12$), Q149A ($n = 21$), R165A ($n = 21$), S168A ($n = 17$), V178A ($n = 19$), M64A/V178A ($n = 15$), V67A/Q80A ($n = 13$), V67A/S168A ($n = 16$), V67A/V178A ($n = 10$), C77A/S168A ($n = 14$), Q80A/R165A ($n = 6$), D95A/R165A ($n = 5$) and Q149A/R165A ($n = 14$) mutants are shown; *n* is the number of biologically independent cells. **c**, Displacement of charge for the isolated VSD in the up, one-click down and two-click down conformations at different transmembrane potentials. Data are mean ± s.d. calculated using the last 40-ns snapshots ($n = 4{,}000$) of 50-ns trajectories. Each system was simulated once at each chosen potential. The gating charge was then calculated as the offset constant between the linear fits, resulting in a gating charge of 1.02 *e* and 0.55 *e* between the up and one-click down, and one-click down and two-click down states, respectively. **d**, Mapping of double-mutant cycle constraints onto up VSD structure. Thick red and green lines connect Cα carbons of interacting pairs. Thin grey lines connect negative-control pairs. **e**, Mapping of literature KAT1 down-state interacting pairs[21] onto up structure. Thick red lines connect Cα carbons of interacting pairs.



Extended Data Fig. 9 | A cysteine–Cd^{2+}–cysteine bridge in the KAT1 VSD promotes channel opening. a, Raw current traces for all four combinations of C77(S) and R165(C). On washing with 100 μM CdCl$_2$, current increases only in the C77/R165C condition (red box, middle), and then decreases again upon EDTA wash. Representative data are shown from the same oocyte, and each experiment was repeated five independent times (five biologically independent oocytes) with similar results. **b,** Pulse protocol used during experiment. **c,** Mapping of C77 (on S1) and R165 (on S4) onto the up VSD structure of KAT1. Cα atoms are indicated by a red line.

Extended Data Table 1 | Cryo-EM data collection, refinement and validation statistics



	KAT1em TMD (EMDB-21018) (PDB 6V1X)	KAT1em Full (EMDB-21019) (PDB 6V1Y)
Data collection and processing		
Magnification	130,000	130,000
Voltage (kV)	300	300
Electron exposure (e–/Å^2)	50	50
Defocus range (μm)	-1 to -2.5	-1 to -2.5
Pixel size (Å)	0.532	0.532
Symmetry imposed	C4	C4
Initial particle images (no.)	124,211	124,211
Final particle images (no.)	91689	91689
Map resolution (Å)	3.5	3.8
FSC threshold	0.143	0.143
Map resolution range (Å)	Not determined	~3.5- 4.5 (ResMap)
Refinement		
Initial model used (PDB code)	*de novo*	*de novo*
Model resolution (Å)	3.5	3.71
FSC threshold	0.143	0.143
Model resolution range (Å)	n/a	n/a
Map sharpening *B* factor (Å^2)	-134	-137
Model composition		
Non-hydrogen atoms	13996	27392
Protein residues	1784	3568
Ligands	8	16
B factors (Å^2)		
Protein	49.0	49.0
Ligand	9.2	9.2
R.m.s. deviations		
Bond lengths (Å)	0.0104	0.0104
Bond angles (°)	1.35	1.35
Validation		
MolProbity score	1.70	1.70
Clashscore	3.64	3.68
Poor rotamers (%)	0	0
Ramachandran plot		
Favored (%)	90.27	90.27
Allowed (%)	9.50	9.50
Disallowed (%)	0.23	0.23

nature research

Corresponding author(s): Eduardo Perozo
Last updated by author(s): Feb 2, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	CryoEM data collection: Gatan Latitude Electrophysiological data collection: Clampex 10.3
Data analysis	CryoEM data analysis: motioncor2, Ctffind4, Relion2, ResMap Model building and analysis: Swiss-model web server, Chainsaw, Coot, Phenix, UCSF Chimera, Chimera X, Segger, MOLE, Hole Molecular Dynamics simulations: VMD, NAMD Electrophysiological data analysis: Clampfit 10.3, in-house python scripts

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Cryo-EM density maps of KAT1 have been deposited in the Electron Microscopy Data Bank under accession codes EMDB-XXXX (full molecule) and EMDB-YYYY (transmembrane-focused refinement). The atomic models of the KAT1 tetramer and octamer have been deposited in the Protein Data Bank under accession code ZZZZ and WWWW, respectively. All other data are available upon reasonable request to the corresponding author.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample size. For cryoEM experiments, particle number was chosen to maximize map quality and resolution. For electrophysiology experiments, cell number was chosen based on convention in the field (at least 4). This was deemed to be sufficient to determine mean and standard error of the mean, allowing for comparison between different mutants.
Data exclusions	For cryoEM experiments, particles were excluded if they did not improve map quality. This is standard practice for cryoEM structure determination. For electrophysiological experiments, recordings were excluded from analysis if leak or endogenous currents prevented analysis. A record was determined to be an outlier, and thus excluded, if the Vh was more than 10 mV (approximately two standard deviations) outside the mean of the normalized ensemble, or if the z was more than two standard deviations outside the mean of the normalized ensemble. This is standard practice in electrophysiology.
Replication	Structure determination was completed once, as is standard. All electrophysiological results contain data from multiple cells, ensuring reproducibility. For mutants that failed to yield currents, at least 10 cells were measured, and the results were confirmed in a separate session of injection and recording. This is standard practice in electrophysiology.
Randomization	Randomization was not employed, as is standard for structural and electrophysiological work
Blinding	Blinding was not employed, as is standard for structural and electrophysiological work

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study	
☒	☐	Antibodies
☐	☒	Eukaryotic cell lines
☒	☐	Palaeontology
☐	☒	Animals and other organisms
☒	☐	Human research participants
☒	☐	Clinical data

Methods

n/a	Involved in the study	
☒	☐	ChIP-seq
☒	☐	Flow cytometry
☒	☐	MRI-based neuroimaging

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Sf9 (ATCC CRL-1711)
Authentication	none
Mycoplasma contamination	not tested
Commonly misidentified lines (See ICLAC register)	not used

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Xenopus oocytes

Wild animals	not used
Field-collected samples	not used
Ethics oversight	University of Chicago Institutional Animal Care and Use Committee, animal usage protocol 71475

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Structural transitions in influenza haemagglutinin at membrane fusion pH

https://doi.org/10.1038/s41586-020-2333-6

Received: 19 December 2019

Accepted: 16 March 2020

Donald J. Benton[1]✉, Steven J. Gamblin[1], Peter B. Rosenthal[2]✉ & John J. Skehel[1]

Infection by enveloped viruses involves fusion of their lipid envelopes with cellular membranes to release the viral genome into cells. For HIV, Ebola, influenza and numerous other viruses, envelope glycoproteins bind the infecting virion to cell-surface receptors and mediate membrane fusion. In the case of influenza, the receptor-binding glycoprotein is the haemagglutinin (HA), and following receptor-mediated uptake of the bound virus by endocytosis[1], it is the HA that mediates fusion of the virus envelope with the membrane of the endosome[2]. Each subunit of the trimeric HA consists of two disulfide-linked polypeptides, HA1 and HA2. The larger, virus-membrane-distal, HA1 mediates receptor binding; the smaller, membrane-proximal, HA2 anchors HA in the envelope and contains the fusion peptide, a region that is directly involved in membrane interaction[3]. The low pH of endosomes activates fusion by facilitating irreversible conformational changes in the glycoprotein. The structures of the initial HA at neutral pH and the final HA at fusion pH have been investigated by electron microscopy[4,5] and X-ray crystallography[6–8]. Here, to further study the process of fusion, we incubate HA for different times at pH 5.0 and directly image structural changes using single-particle cryo-electron microscopy. We describe three distinct, previously undescribed forms of HA, most notably a 150 Å-long triple-helical coil of HA2, which may bridge between the viral and endosomal membranes. Comparison of these structures reveals concerted conformational rearrangements through which the HA mediates membrane fusion.

We have used single-particle cryo-electron microscopy (cryo-EM) of the HA ectodomain to obtain structural information on the transitions between HAs at neutral pH and fusion pH. Specifically, we have examined the structures formed on incubation of HA at pH 5.0 and 4 °C as a function of time, using incubation times of 10 s, 20 s, 60 s and 30 min. We have characterized three structural intermediates that have been trapped through rapid freezing (by plunging grids into liquid ethane).

For each incubation time we image and classify a number of different conformations of HA (Fig. 1). The ratio of these species changes during the time course (Fig. 1b). We are able to describe three distinct, previously undescribed forms of HA (states II, III and IV) in addition to the neutral-pH (state I) and fusion-pH (state V) structures[6,7] (Fig. 1). We interpret the new forms as sequential intermediate states between the two previously known conformations. At the earliest sampling time of 10 s, we observe dilated HA structures in which the envelope-distal domains tilt away from the axis of the HA trimer and the membrane-proximal domain is disordered (state II). At 20 s another population is more evident, in which the membrane-distal domain is more dilated and the membrane-proximal region shows disorder and rearrangement (state III). At 20 s and 60 s we observe, increasingly, the most notable of the structural changes: a 150 Å-long α-helical coil projecting between the dilated membrane-distal domains (state IV).

Dilation of the membrane-distal domains

In state II, the globular HA1 domain rotates and alters the interactions between the subunits of the trimer (Fig. 2a). The rotation is small, increasing the distance between centroids of the domains to 38 Å, from 35 Å in the neutral-pH form (Extended Data Fig. 1a).

The requirement for these domains to separate for membrane fusion was previously concluded from experiments in which the low-pH-dependent conformational change was blocked by crosslinking the domains, either by introducing disulfide bonds[9] or by antibody binding[10]. There are also reports of similar subunit dissociation in antigenic analyses of HAs at neutral pH, as judged by accessibility of the intersubunit interface to specific antibodies[11,12]. In our experiments, which were done at 4 °C, domain dilation was seen only at pH 5.0 and not in a parallel experiment at pH 8.0.

In state II and the other states, the membrane-distal parts of the dilated domains (HA1 residues 45–310) as monomers are indistinguishable from the equivalent regions of the neutral-pH form. The stability of this domain in isolation is consistent with X-ray crystallographic observations of the membrane-distal domain isolated from HA in the fusion-pH conformation[13]. As well as the rearrangements in the HA1 membrane-distal domain, there is also an indication of disorder in the membrane-proximal regions of HA, as shown by local-resolution estimates (Fig. 2b and Extended Data Fig. 2). There is, however, little evidence of large structural rearrangements.

---

[1]Structural Biology of Disease Processes Laboratory, Francis Crick Institute, London, UK. [2]Structural Biology of Cells and Viruses Laboratory, Francis Crick Institute, London, UK. ✉e-mail: donald.benton@crick.ac.uk; peter.rosenthal@crick.ac.uk

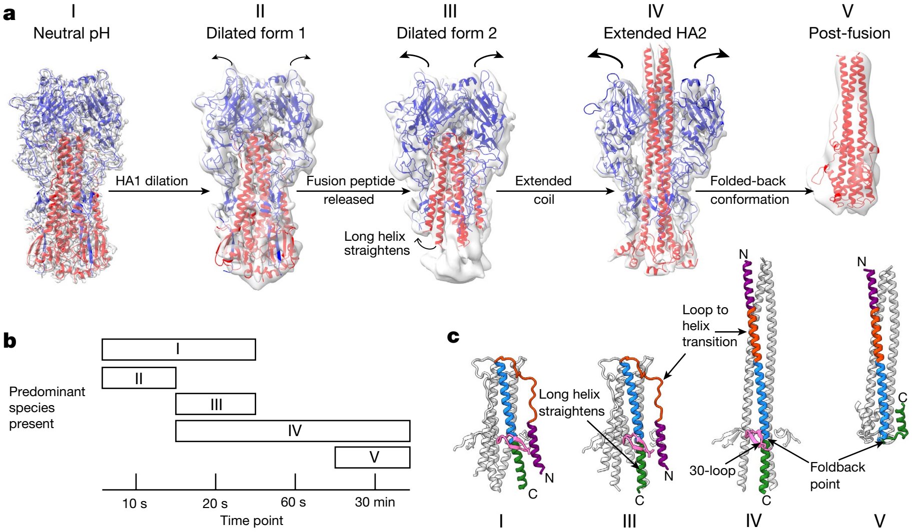


Fig. 1 | HA fusion intermediates. a, Cryo-EM reconstructions of HA at fusion pH from various time points after acidification. Cryo-EM maps are shown (grey) with models of HA1 (blue) and HA2 (red). Structures for states I (indistinguishable from the neutral-pH state), III and IV are from 20 s, and for state II from 10 s. State V was obtained with a 30-min incubation, supplemented with 2-mercaptoethanol, dissociating the disulfide-linked HA1. The state V model shown is the previously determined crystal structure of fusion-pH HA2 (Protein Data Bank (PDB, https://www.ebi.ac.uk/pdbe/) code 1HTM; ref. [7]). **b**, Distribution of different species at chosen time points. **c**, Rearrangements in HA2 residues 38–125 associated with conformational states I, III, IV and V. The 30-loop (HA1 residues 22–37) is in pink; HA2 is coloured by residue number: purple, 38–55; orange, 56–75; blue, 76–105; green, 106–125.

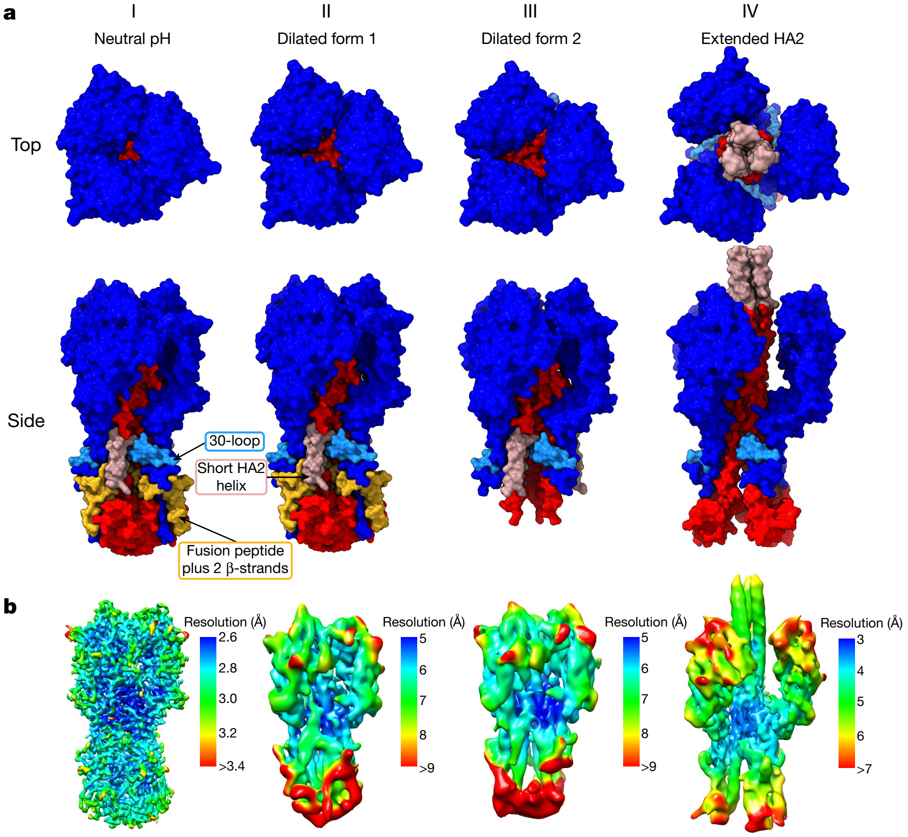


Fig. 2 | Surface representations of HA fusion intermediates. a, Molecular surfaces for states I–IV. HA1 is coloured in dark blue and the 30-loop (residues 22–37) in light blue. HA2 is in red, with residues 1–37 (containing the fusion peptide, residues 1–23) in yellow and the short helix (residues 38–55) in pink. Top views of surfaces show the increasing dilation of the membrane-distal domains. Side views show several features. Between states II and III, the fusion peptide and attached two β-strands are absent while the base of the protein becomes disordered. Between states III and IV, the 150 Å coiled-coil forms between the dilated HA1 domains. The base of HA2 has also opened up in state IV compared with state I. **b**, Cryo-EM maps, coloured by local-resolution estimations (contour bars are state-specific).

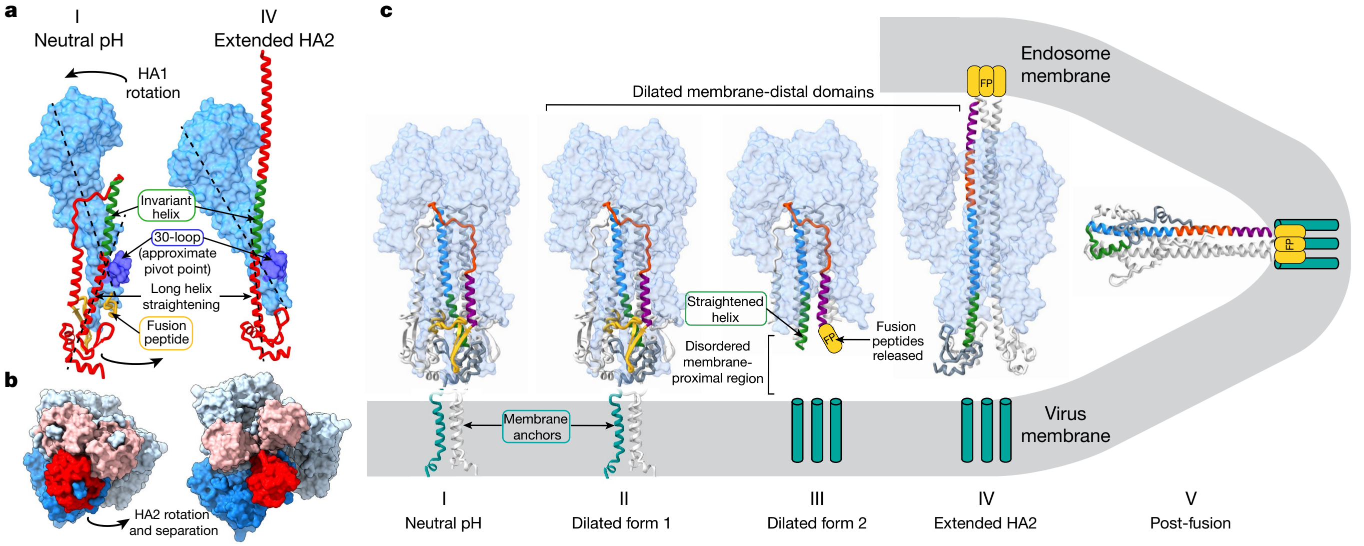


Fig. 3 | Structural rearrangements of HA fusion intermediates. a, Concerted rearrangements of HA1 and HA2 between states I and IV. HA1 is shown as a molecular surface in light blue, with the 30-loop in dark blue. HA2 is shown as a red ribbon with the invariant helix in green. The fusion peptide in state I is coloured yellow, but not ordered in state IV. Dotted lines show the trajectory of the membrane-proximal region of the extended helix in one case and the approximate long axis of HA1 in the other. HA1 rotates as a rigid body as the long helix of HA2 straightens relative to state I. This concerted motion is transmitted by the attachment of HA1 to HA2 via the 30-loop, which is the approximate pivot point of the HA1 rotation. The HA2 helix straightens into the space previously occupied by the now-displaced fusion peptide. **b**, Orthogonal view to **a**; membrane-proximal regions of HA2 open upon this concerted motion, transitioning from a closely packed neutral-pH conformation (state I) to the opened extended HA2 form (state IV). Images are molecular surfaces, with the monomers coloured as in **a**. **c**, Diagram showing HA membrane-fusion intermediates. HA1 is shown as a molecular surface with HA2 as a ribbon. One HA2 monomer is coloured by sequence: yellow, residues 1–37; purple, 38–55; orange, 56–75; blue, 76–105; green, 106–125; grey, 125–175. The location of the fusion peptide, not resolved in states III–V, is shown schematically (yellow). Membrane anchors (cyan), not present in the protein used here, are represented as blocks in the virus and endosome membranes in states III and IV. For states I and II the membrane anchor is that of another HA subtype, H1 (ref. [22]). State IV is depicted as interacting with the virus membrane (via membrane anchors) and the endosomal membrane (via the fusion peptide) before refolding to state V, where the fusion peptide and membrane anchors are shown as colocated.

State III shows further dilation of the membrane-distal domain (Fig. 2a) and a substantial rearrangement in the membrane-proximal region of HA2. The further dilation increases the average distance between centroids of the domain to 40 Å from 38 Å in state II (Extended Data Fig. 1a). Both states II and III appear to retain an interaction between the H3 subtype-specific N-linked glycan at Asn165 and Trp222 (Extended Data Fig. 1b).

The membrane-proximal regions of this state are disordered, as indicated by the low local resolution (Fig. 2b), which involves a loss of density for the fusion peptide (residues 1–23) and for the two attached membrane-proximal β-strands (residues 24–38). The shorter α-helix of HA2 (residues 38–59) remains intact, but is displaced towards a neighbouring longer α-helix. The loss of density for the fusion peptide is accompanied by a rearrangement of the carboxy (C) terminus of the long α-helix of HA2, in which residues 100–125 straighten into the space previously occupied by the fusion peptide (similar to state IV, Fig. 3a). States II and III both have an extension of density at the C terminus of the shorter α-helix of HA2, indicating elongation of the α-helix by inclusion of residues from the interhelical loop (Extended Data Fig. 1c).

Extended coil conformation

State IV contains a 150 Å trimeric coil formed from a single continuous helix of residues 38–125 of HA2, which includes the short α-helix, the interhelical loop and the long α-helix of HA2 present in the neutral-pH structure (Fig. 1c). States I–IV share the common section of HA2 (residues 76–98), which forms a coiled-coil, extended in state IV from its amino (N) terminus by the interhelical loop (residues 59–75) and the shorter α-helix (residues 38–58) of HA2. The C-terminal region of the coil is formed by the remainder of the long central α-helix (residues 99–125). The coil projects between the membrane-distal domains (Figs. 1, 2), which are dilated further, creating additional space around the trimer axis and increasing the displacement between the membrane-distal domains to 48 Å, from 40 Å in state III (Extended Data Fig. 1a).

The membrane-distal domains in state IV exhibit lower local resolution (Fig. 2b), indicating structural heterogeneity, which we examined using asymmetric classifications (Extended Data Fig. 3). We obtained the state IV structure by applying threefold symmetry. The heterogeneity in the positions of these domains limits the global resolution to 4.0 Å and this can be improved to 3.6 Å, with some anisotropy, by subtracting membrane-distal regions from the particles.

HA2 residues 1–37—which include the fusion peptide and two strands of the five-stranded membrane-proximal β-sheet—are not detected in state IV, presumably indicating disorder with respect to the symmetric coiled-coil. The structure of the coiled-coil component of the 150 Å helix (residues 38–106) is very similar to those of the coiled-coils reported for the 110 Å structure of the HA-derived fusion-pH fragment[7] and the fusion-pH conformation of *Escherichia coli*-expressed HA2 (ref. [8]) determined by X-ray crystallography. They all extend to residue 38, which is a component of the helical cap structure in the *E. coli*-expressed molecule.

The formation of the extended structure described here is a reflection of the propensity of the residues involved to adopt a helical conformation. A similar proposal was made for a 150 Å α-helical coil in neutral-pH HA on the basis of the amino-acid sequence of HA2 (ref. [14]), and for fusion-pH HA from studies of synthetic peptides[15]. Both predictions were made before the crystal structures revealed that the relevant forms adopt hairpin structures. The extended intermediate described here indicates the stability of such a helical structure before hairpin formation occurs. It is also consistent with cryotomography showing influenza virus HA interacting with target liposomes at low pH[16].

The rearrangement of the remaining C-terminal regions of HA2 (residues 126–169) opens the structure by an outward rigid body rotation that accompanies the straightening of the coil (Fig. 3b). However, the density for this region is poorly defined (Fig. 2b), indicating a higher degree of mobility of the domains, as they lose the trimeric contacts observed at neutral pH.

The 30-loop

Insights into the concerted nature of HA1 and HA2 rearrangements are revealed by the structure and function of the 30-loop (HA1 residues 22–37). States I, II, III and IV all retain density for this loop, which is inserted in the region of residues 104–107 of HA2, in the long central α-helices, about 50 Å from the virus membrane (Fig. 1c).

The location of the interaction of the 30-loop with the long helix of HA2 is similar in states I and IV. There are, however, several changes in the side chains contacted in state IV that are associated with the relocation of the short helix of HA2. In this state, the 30-loop makes interactions with His106 and Gln105 at the site of the 180° turn in the fusion-pH structure (Extended Data Fig. 4).

There are probably two roles for the 30-loop in the refolding process. First, it may couple changes in the orientation of the membrane-distal domains with helical rearrangements in the membrane-proximal domain. The interactions of this loop with HA2 occur at the approximate pivot point of HA1, and connect the dilation of the membrane-distal domains with the straightening of the HA2 helix that accompanies fusion-peptide release in states III and IV. Second, the 30-loop is positioned to stabilize the extended coil in state IV. The high conservation and importance of these interactions for membrane fusion has been noted in studies of mutant HAs that differ in their stability[17–20], but, until now, without any obvious mechanism (Extended Data Fig. 4).

Concerted structural changes for fusion

This cryo-EM study and earlier X-ray crystallographic analyses[7,8] indicate that the N-terminal part of HA2, up to residue 38, is flexible. Structure has been assigned to an analogue of the fusion peptide (residues 1–23) by nuclear magnetic resonance (NMR) spectroscopy[21]. Taken together, these results imply a flexibility of the region linking the α-helical coil to the fusion peptide (residues 24–37) that may be a requirement for the formation of the fusion-pH structure. We came to a similar conclusion from our previous cryo-EM analysis[22] of a flexible region that links the membrane-anchor α-helices of neutral-pH, full-length HA to its ectodomain. The flexibility in these N- and C-terminal regions of HA2 would seem to be a requirement for accommodating the large-scale structural rearrangements and the close approach of the two membranes before fusion (Fig. 3c).

Our analysis of different states in the refolding of HA2 at fusion pH suggests staged, concerted changes of HA in its membrane-proximal and membrane-distal regions. The coincidence of these domain rearrangements also suggests that they are relayed throughout the length of the molecule to result in extrusion of the fusion peptide at the terminus of the α-helical coil through the dilated membrane-distal domain, locating the N and C termini of HA2 in opposing membranes. Residues 38–106 in the extended coil are identical to the fusion-pH structures previously obtained, suggesting that the coiled-coil serves as a scaffold for the subsequent refolding of the C-terminal regions to colocate the fusion-peptide and membrane-anchor domains in the same membrane (Fig. 3c).

In describing the three previously unknown structures involved in HA refolding, we have also identified potential functions in membrane fusion for the conserved loop formed by residues 22–37 of HA1, the 30-loop. This loop acts to stabilize the extended coil and delay reversal of the coil through the 180° turn, for which displacement of the 30-loop is necessary. This could enable effective interaction of the fusion peptide with the endosomal membrane, before the membrane-anchor region and the fusion peptide become positioned together at one end of the fusion-pH structure and in fused membranes (Fig. 3c). As a consequence, the interactions made by the 30-loop may be required for the controlled delivery of the membrane anchor and the virus membrane to the site of fusion.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2333-6.

1. Cossart, P. & Helenius, A. Endocytosis of viruses and bacteria. *Cold Spring Harb. Perspect. Biol*. **6**, a016972 (2014).
2. Skehel, J. J. & Wiley, D. C. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annu. Rev. Biochem*. **69**, 531–569 (2000).
3. Brunner, J. Testing topological models for the membrane penetration of the fusion peptide of influenza virus hemagglutinin. *FEBS Lett*. **257**, 369–372 (1989).
4. Ruigrok, R. W. et al. Electron microscopy of the low pH structure of influenza virus haemagglutinin. *EMBO J*. **5**, 41–49 (1986).
5. Ruigrok, R. W. H. et al. Studies on the structure of the influenza virus haemagglutinin at the pH of membrane fusion. *J. Gen. Virol*. **69**, 2785–2795 (1988).
6. Wilson, I. A., Skehel, J. J. & Wiley, D. C. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* **289**, 366–373 (1981).
7. Bullough, P. A., Hughson, F. M., Skehel, J. J. & Wiley, D. C. Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature* **371**, 37–43 (1994).
8. Chen, J., Skehel, J. J. & Wiley, D. C. N- and C-terminal residues combine in the fusion-pH influenza hemagglutinin HA_2 subunit to form an N cap that terminates the triple-stranded coiled coil. *Proc. Natl Acad. Sci. USA* **96**, 8967–8972 (1999).
9. Godley, L. et al. Introduction of intersubunit disulfide bonds in the membrane-distal region of the influenza hemagglutinin abolishes membrane fusion activity. *Cell* **68**, 635–645 (1992).
10. Barbey-Martin, C. et al. An antibody that prevents the hemagglutinin low pH fusogenic transition. *Virology* **294**, 70–74 (2002).
11. Watanabe, A. et al. Antibodies to a conserved influenza head interface epitope protect by an IgG subtype-dependent mechanism. *Cell* **177**, 1124–1135.e16 (2019).
12. Bangaru, S. et al. A site of vulnerability on the influenza virus hemagglutinin head domain trimer interface. *Cell* **177**, 1136–1152 (2019).
13. Bizebard, T. et al. Structure of influenza virus haemagglutinin complexed with a neutralizing antibody. *Nature* **376**, 92–94 (1995).
14. Ward, C. W. & Dopheide, T. A. Influenza virus haemagglutinin. Structural predictions suggest that the fibrillar appearance is due to the presence of a coiled-coil. *Aust. J. Biol. Sci*. **33**, 441–447 (1980).
15. Carr, C. M. & Kim, P. S. A spring-loaded mechanism for the conformational change of influenza hemagglutinin. *Cell* **73**, 823–832 (1993).
16. Calder, L. J. & Rosenthal, P. B. Cryomicroscopy provides structural snapshots of influenza virus membrane fusion. *Nat. Struct. Mol. Biol*. **23**, 853–858 (2016).
17. Daniels, R. S. et al. Fusion mutants of the influenza virus hemagglutinin glycoprotein. *Cell* **40**, 431–439 (1985).
18. Wharton, S. A., Skehel, J. J. & Wiley, D. C. Studies of influenza haemagglutinin-mediated membrane fusion. *Virology* **149**, 27–35 (1986).
19. Xu, R. & Wilson, I. A. Structural characterization of an early fusion intermediate of influenza virus hemagglutinin. *J. Virol*. **85**, 5172–5182 (2011).
20. Collins, P. J. et al. Recent evolution of equine influenza and the origin of canine influenza. *Proc. Natl Acad. Sci. USA* **111**, 11175–11180 (2014).
21. Lorieau, J. L., Louis, J. M., Schwieters, C. D. & Bax, A. pH-triggered, activated-state conformations of the influenza hemagglutinin fusion peptide revealed by NMR. *Proc. Natl Acad. Sci. USA* **109**, 19994–19999 (2012).
22. Benton, D. J. et al. Influenza hemagglutinin membrane anchor. *Proc. Natl Acad. Sci. USA* **115**, 10112–10117 (2018).

Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Protein preparation

The HA ectodomain was isolated from purified X-31(H3N2) virus, which contains the HA from A/Aichi/1/68. Virus was propagated in 11- to 12-day-old hens' eggs and incubated for 48 h. Virus was purified from allantoic fluid by sucrose gradient centrifugation. HA ectodomain was released from the virus by detergent extraction with 2% (w/v) octyl-β-glucoside, followed by overnight digestion at room temperature with 5% (w/w) trypsin, which produces a fragment with the membrane anchor removed (HA1, and HA2 residues 1–174). Protein was further purified by anion-exchange chromatography and size-exclusion chromatography (SDS–PAGE shown in Extended Data Fig. 5). The final protein buffer was 25 mM Tris pH 8.0, 150 mM NaCl.

Cryo-EM sample preparation and data collection

Low-pH-induced conformations of HA were obtained by mixing with low-pH buffer followed by an incubation at 4 °C and plunge freezing. We mixed 2.5 mg ml^{-1} HA 1:1 with 0.1 M citrate pH 4.9 plus 0.1% octyl-β-glucoside, which gave a final pH of 5.0. Octyl-β-glucoside was added to reduce orientational bias, as used previously[22]. The mixture was rapidly applied to a grid equilibrated to 4 °C in 100% humidity for a set time period to obtain pH incubation times of 10 s, 20 s and 60 s. The 30-min time point was prepared by mixing HA with the low-pH buffer as above, supplemented with 10 mM 2-mercaptoethanol, followed by a 30-min incubation on ice. Specimens of HA at pH 8.0 were prepared by plunge freezing HA at 2.5 mg ml^{-1} concentration supplemented with 0.1% octyl-β-glucoside. All samples were prepared by applying 4 µl of sample to an R2/2 300 mesh Quantifoil grid, followed by a 4–4.5 s blot using a Vitrobot MkIII, plunge freezing into liquid ethane.

Data were collected using a Titan Krios microscope operating at 300 kV. Micrographs were collected using a Falcon 3 detector in electron-counting mode. Exposures were 60 s with a total dose of 33.9 e $Å^{-2}$, fractionated into 30 frames, with a calibrated pixel size of 1.09 Å. Images were collected using a defocus of 1.5–3 µm.

Data processing

Movie frames were aligned using MotionCor2 (ref. [23]) and contrast transfer function was fitted using CTFFIND4 (ref. [24]). All data processing was carried out using both RELION[25] and cryoSPARC[26]. Particles were picked using crYOLO[27] by training models on manually picked micrographs. Data-processing workflows for each time point are shown in Extended Data Figs. 6–9.

For 10 s, 20 s and 60 s time points, particles were subjected to two rounds of Relion two-dimensional (2D) classification, retaining classes that exhibited clear secondary-structure features (representative classes are shown in Extended Data Fig. 5). Initial models were generated using a cryoSPARC ab initio reconstruction initiated with two classes for the 20 s data. This generated models for the neutral-pH-like and extended-intermediate-like conformations. These two maps were used as initial models for a two-class Relion three-dimensional (3D) classification to sort particles into neutral-pH-like and extended-intermediate-like classes for each of the time points.

For the 20 s data set, the neutral pH-like conformation particles were further separated using RELION 3D classification, which generated a single class of neutral-pH HA and another of the dilated form 2. The 124,000 particles that made up the dilated form 2 were refined using cryoSPARC, yielding a map at 5.6 Å resolution (state III). There were 157,000 particles that made up the neutral-pH conformation. For these particles, we carried out CTF refinement and Bayesian polishing[28,29] using RELION. These particles were then refined using cryoSPARC to obtain a map at 3.0 Å resolution (state I).

The intermediate-like particles were first classified using a two-class heterogeneous refinement in cryoSPARC to separate out a class with a stronger membrane-proximal region of 546,000 particles. These particles were then refined in RELION and Bayesian polishing was carried out. Particles were then refined using cryoSPARC imposing C3 symmetry, generating a map at 4.0 Å resolution. Particle subtraction was carried out using RELION to remove the membrane-distal parts of HA1. Subtracted particles were refined using cryoSPARC, imposing C3 symmetry. This generated a map with a global resolution of 3.6 Å (subtracted state IV). In order to further improve the membrane-proximal density of the extended-intermediate particles, particle subtraction was carried out to leave only the density of this region. A three-class 3D classification without alignments was carried out, which generated a class of 189,000 particles with a stronger membrane-proximal density. The unsubtracted particles were then refined using cryoSPARC homogeneous refinement imposing C3 symmetry to generate a map with a global resolution of 4.2 Å (state IV).

To examine the flexibility of the HA1 regions of extended-intermediate particles, which had been refined imposing C3 symmetry, signal subtraction was carried out to remove membrane-proximal regions. These particles were then classified into ten classes with a 3D classification using RELION-relax[30], allowing asymmetric classification with C3 symmetry priors with a limited angular search range. The original unsubtracted particles from each class were then refined asymmetrically using cryoSPARC homogeneous refinement.

For the 10 s data, the neutral-pH-like particles were further separated using RELION 3D classification, yielding a single class, which resembled the neutral-pH form of HA (105,000 particles) and another which made up the dilated form 1 (72,000 particles). These dilated form 1 particles were refined using cryoSPARC to obtain a map at 5.5 Å resolution (state II).

Particles from the 30-min plus 2-mercaptoethanol time point were picked, using crYOLO, generating 2.24 million particles. 2D classification was carried out using RELION. Classes with particles that resembled the post-fusion form of HA were selected (359,000) and subjected to a two-class cryoSPARC heterogeneous refinement using an initial model generated using the cryoSPARC ab initio reconstruction. The best class of 184,000 particles was refined using RELION, imposing C3 symmetry. The refined particles were classified using a 3D classification without particle alignments. The best class (of 62,000 particles) was chosen and refined using RELION, imposing C3 symmetry, generating a map at 9.9 Å resolution. We do not characterize other structural intermediates, other than the states listed above, possibly because of their low population or heterogeneity.

For the neutral-pH HA condition, data were collected at pH 8.0; 2,819 micrographs yielded 643,000 particles. After two rounds of 2D classification, 320,000 particles remained. These were 3D classified with RELION. The most-populated class of 237,000 particles had CTF refinement and Bayesian polishing carried out using RELION. The particles were further 3D classified without alignment, following refinement, which gave the most-populated class of 200,000 particles. These particles were refined using cryoSPARC to give a final map with a global resolution of 3.0 Å, which was indistinguishable from the neutral-pH-like structure obtained at pH 5.0 (state I).

Model building

Before model building, we determined the local resolution of maps using blocres[31] implemented in cryoSPARC. Maps were automatically sharpened[32] and locally filtered using cryoSPARC. Models were built using parts of previously determined structures of X-31 HA (PDB identification code 2YPG)[33] and the post-fusion conformation from *E. coli*-expressed HA2 (1QU1)[8]. Models were manually adjusted using

COOT[34], with real-space refinement and validation carried out using PHENIX[35]. Measurements of domain displacement and figures were made using Chimera[36].

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Maps and models have been deposited in the Electron Microscopy Data Bank (http://www.ebi.ac.uk/pdbe/emdb/, under accession numbers EMD-10696, EMD-10697, EMD-10698, EMD-10699, EMD-10700, EMD-10701, and EMD-10702). Models have been deposited in the Protein Data Bank (https://www.ebi.ac.uk/pdbe/, under identification codes 6Y5G, 6Y5H, 6Y5I, 6Y5J, 6Y5K and 6Y5L). The raw image for Extended Data Fig. 5c is provided in Supplementary Fig. 1.

23. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
24. Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
25. Scheres, S. H. W. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
26. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
27. Wagner, T. et al. SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Commun. Biol.* **2**, 218 (2019).
28. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
29. Zivanov, J., Nakane, T. & Scheres, S. H. W. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCrJ* **6**, 5–17 (2019).
30. Ilca, S. L. et al. Multiple liquid crystalline geometries of highly compacted nucleic acid in a dsRNA virus. *Nature* **570**, 252–256 (2019).
31. Cardone, G., Heymann, J. B. & Steven, A. C. One number does not fit all: mapping local variations in resolution in cryo-EM reconstructions. *J. Struct. Biol.* **184**, 226–236 (2013).
32. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).
33. Lin, Y. P. et al. Evolution of the receptor binding properties of the influenza A(H3N2) hemagglutinin. *Proc. Natl Acad. Sci. USA* **109**, 21474–21479 (2012).
34. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
35. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
36. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
37. Rosenthal, P. B. et al. Structure of the haemagglutinin-esterase-fusion glycoprotein of influenza C virus. *Nature* **396**, 92–96 (1998).

Acknowledgements We thank A. Nans of the Structural Biology Science Technology Platform for assistance with data collection; P. Walker and A. Purkiss of the Structural Biology Science Technology Platform; the Scientific Computing Science Technology Platform for computational support; and L. Calder for discussions. This work was funded by the Francis Crick Institute, which receives its core funding from Cancer Research UK (grant numbers FC001078 and FC001143), the UK Medical Research Council (FC001078 and FC001143), and the Wellcome Trust (FC001078 and FC001143).

Author contributions D.J.B carried out research and collected and analysed data. All authors conceived and designed the research and wrote the paper.

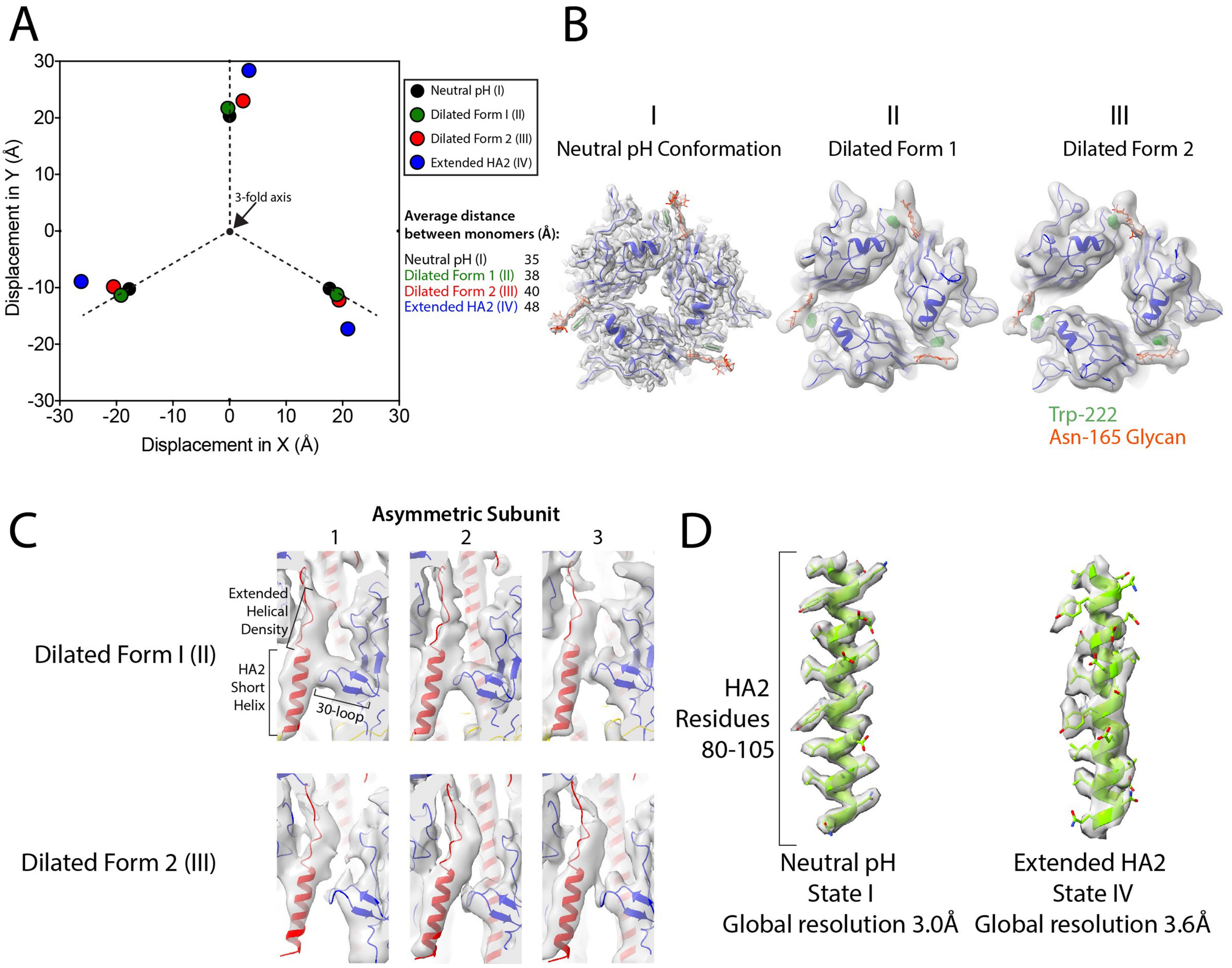
Competing interests The authors declare no competing interests.

Additional information

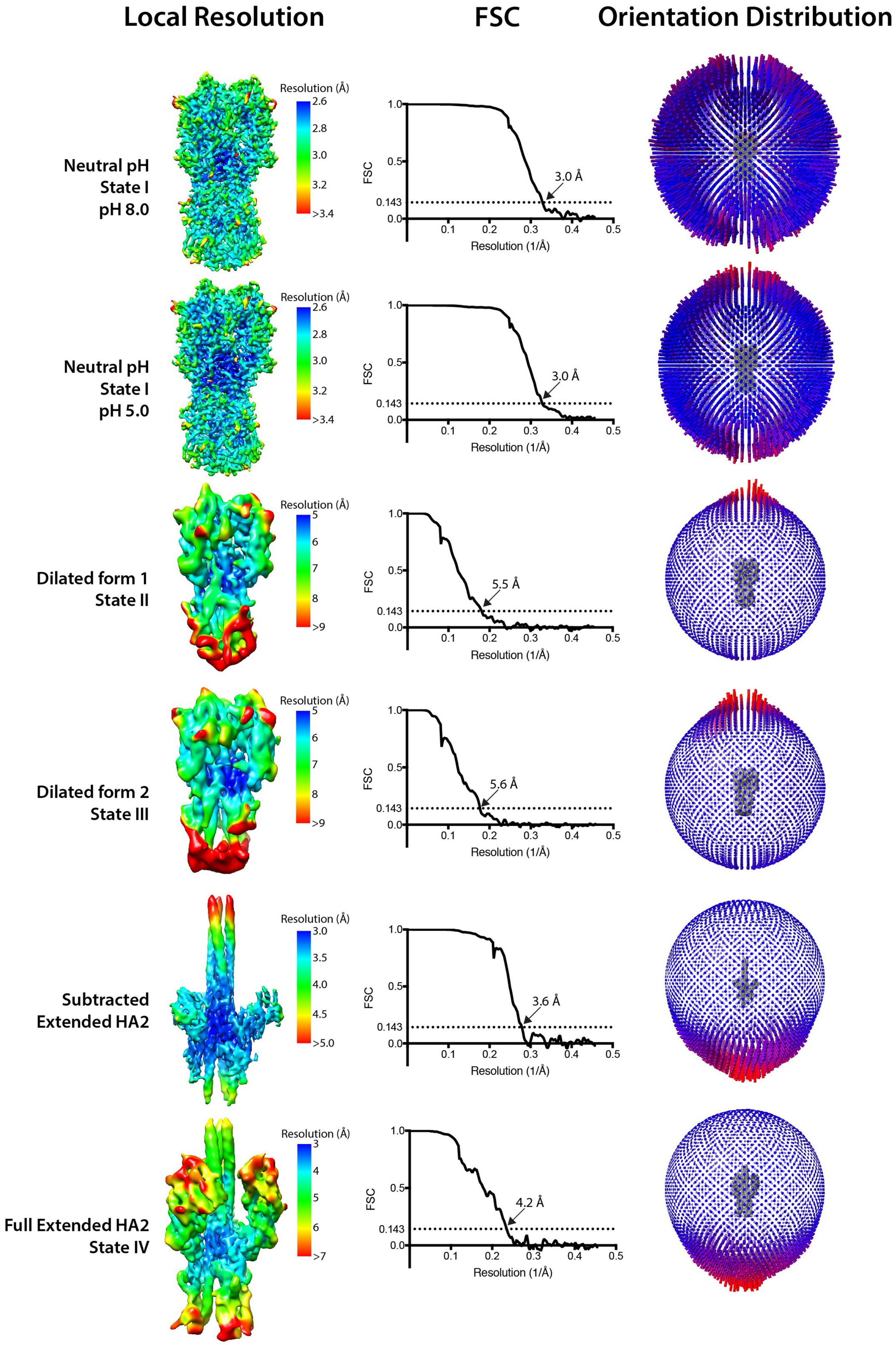
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2333-6.

Correspondence and requests for materials should be addressed to D.J.B. or P.B.R.

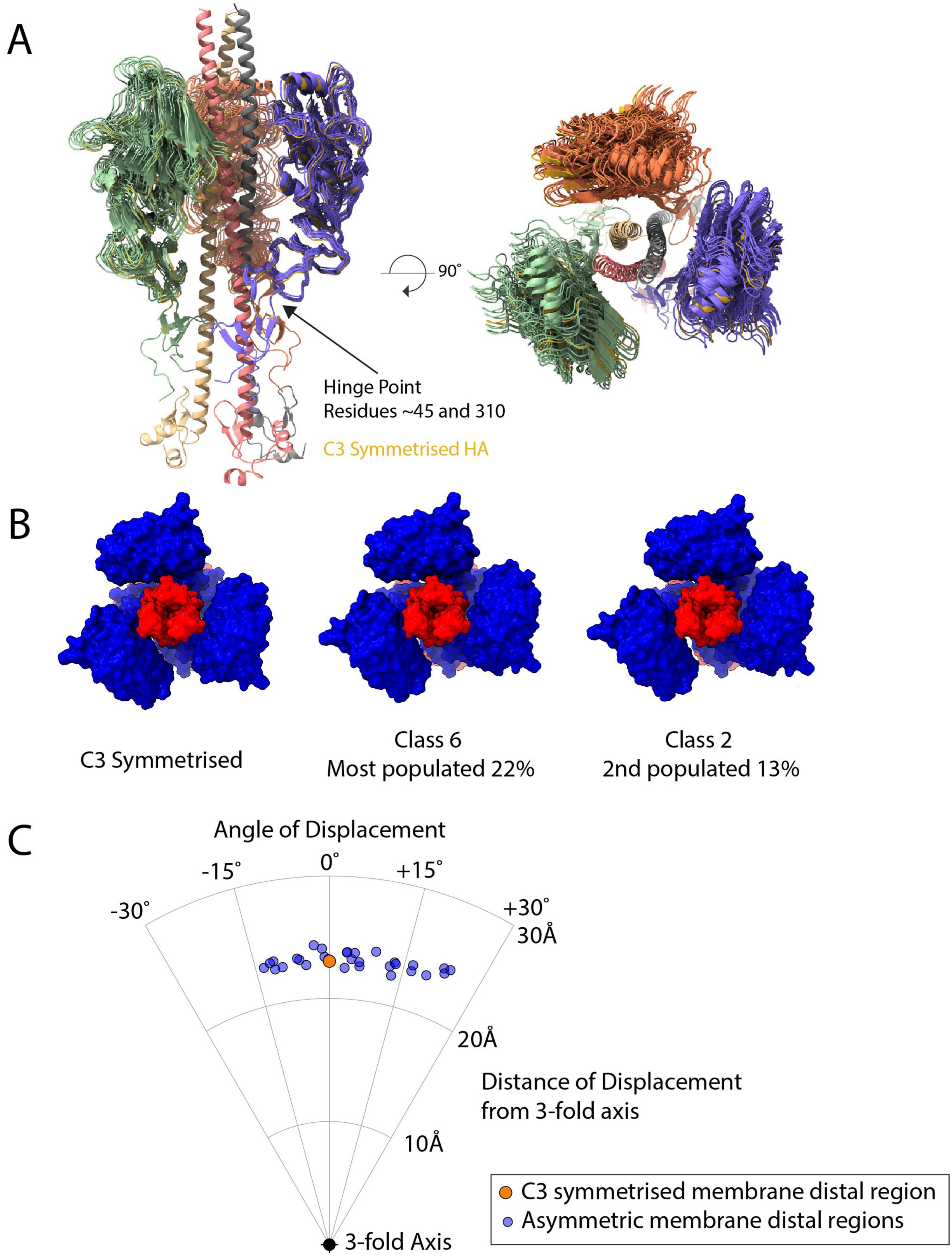
Peer review information *Nature* thanks Stephen Harrison, Juha Huiskonen and Yorgo Modis for their contribution to the peer review of this work.



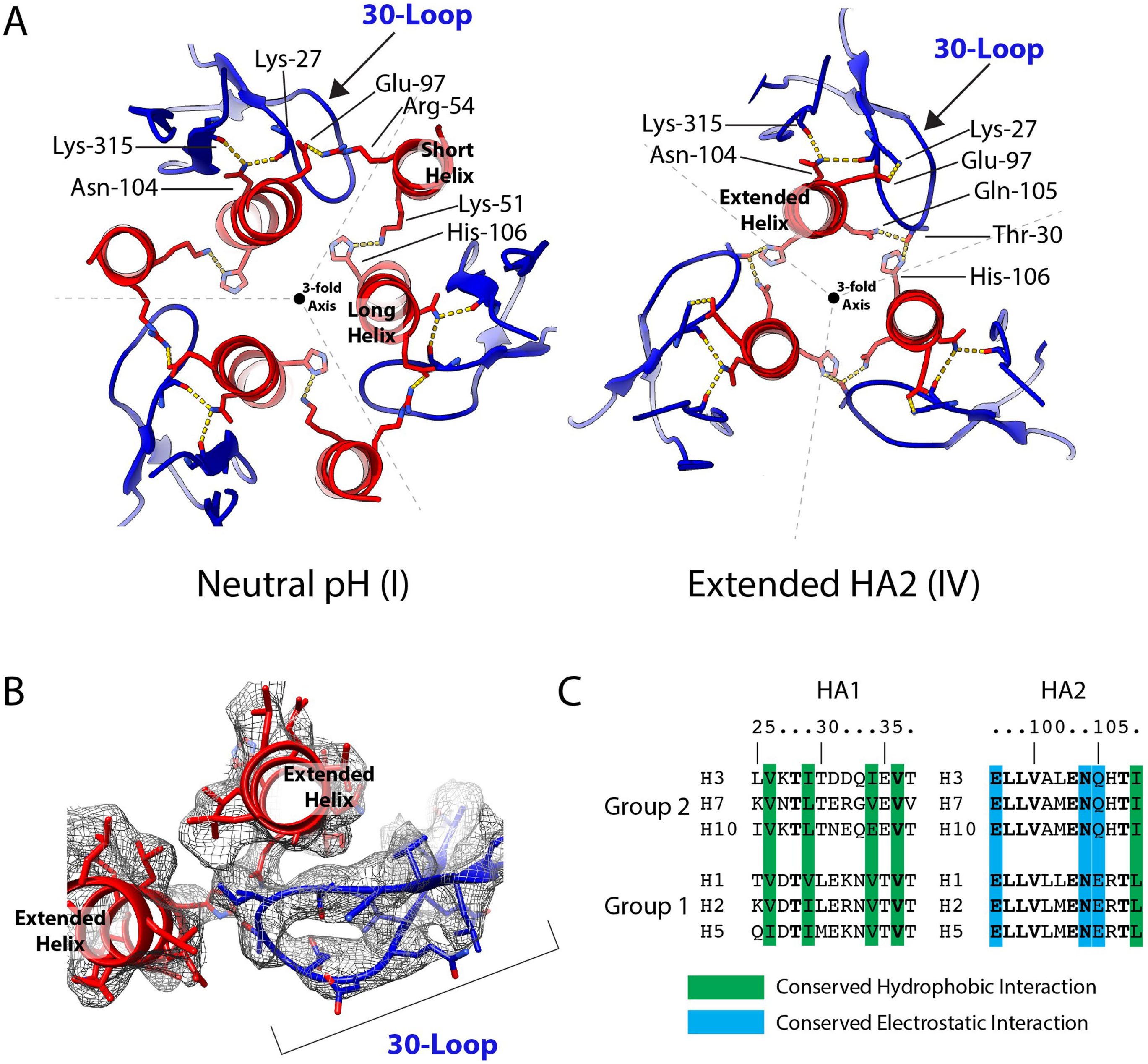
Extended Data Fig. 1 | Features of intermediate states. a, Displacement measurements between dilated HA1 domains. Measurements are the position of the centroid of HA1 residues 45–310, with structures aligned on the invariant parts of HA2 (residues 76–98). **b**, An interaction can be seen between the glycan attached to Asn165 (orange) and Trp222 (green) on the neighbouring HA1 monomer (blue), crosslinking the HA1 domains. This is present in states I, II and III. **c**, Cryo-EM density (grey) and model, with HA1 in blue and HA2 in red, for the short helix of HA2 (residues 38–55) and the 30-loop (HA1 residues 22–37). Density extends beyond the top of the helix, indicating elongation of this helix by residues from the interhelical loop. **d**, Example of typical density (grey) for the same section of an HA2 helix from states I and IV (green).



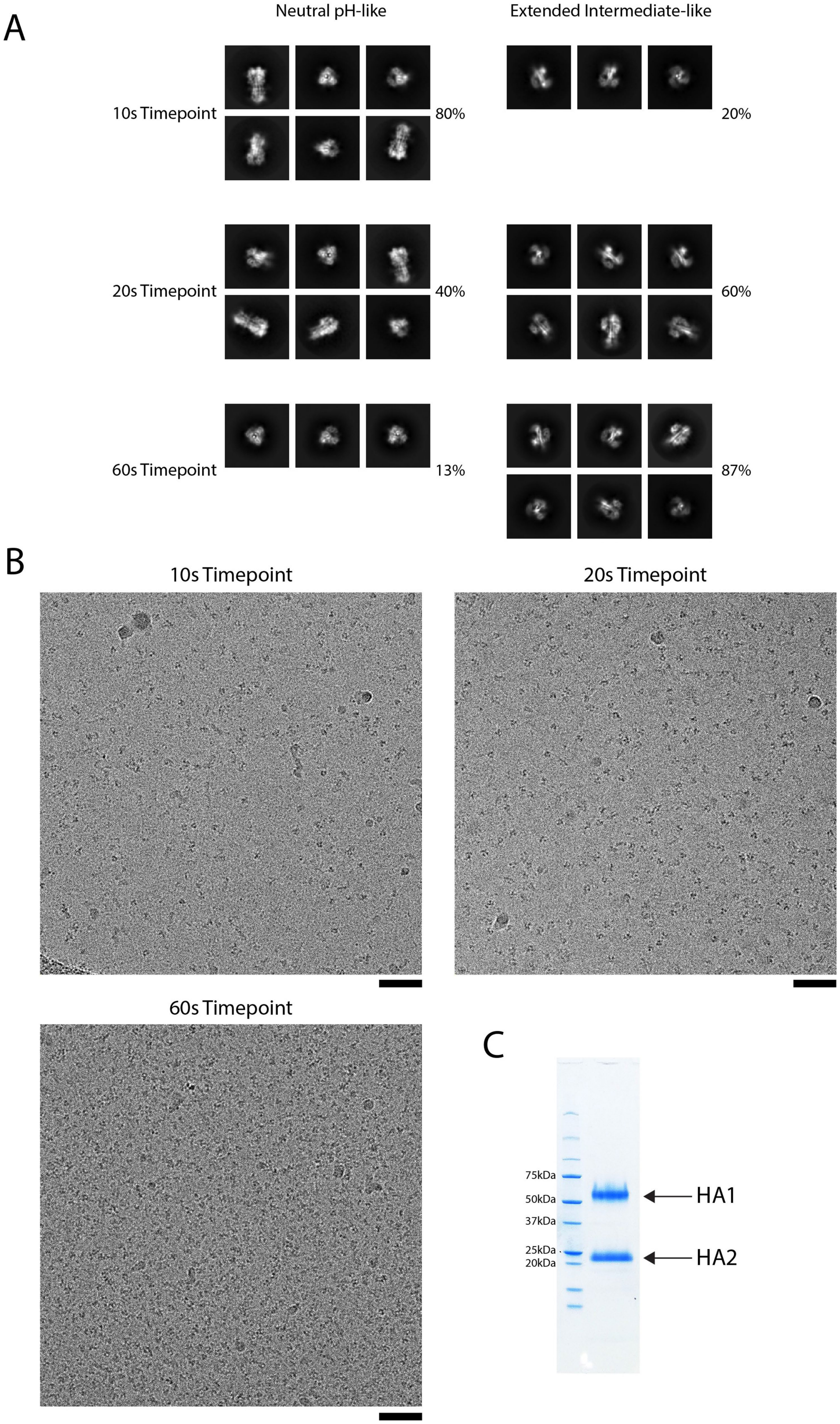
Extended Data Fig. 2 | Resolution estimates of obtained cryo-EM maps. Local-resolution estimates, global-resolution FSC curves and orientation-distribution plots for different conformations of HA. FSC curves have values for FSC of 0.143, shown as the resolution criterion. Reconstruction of state V has a resolution of 9.9 Å and has little variability in local resolution. See Extended Data Fig. 9 for the FSC curve for state V.



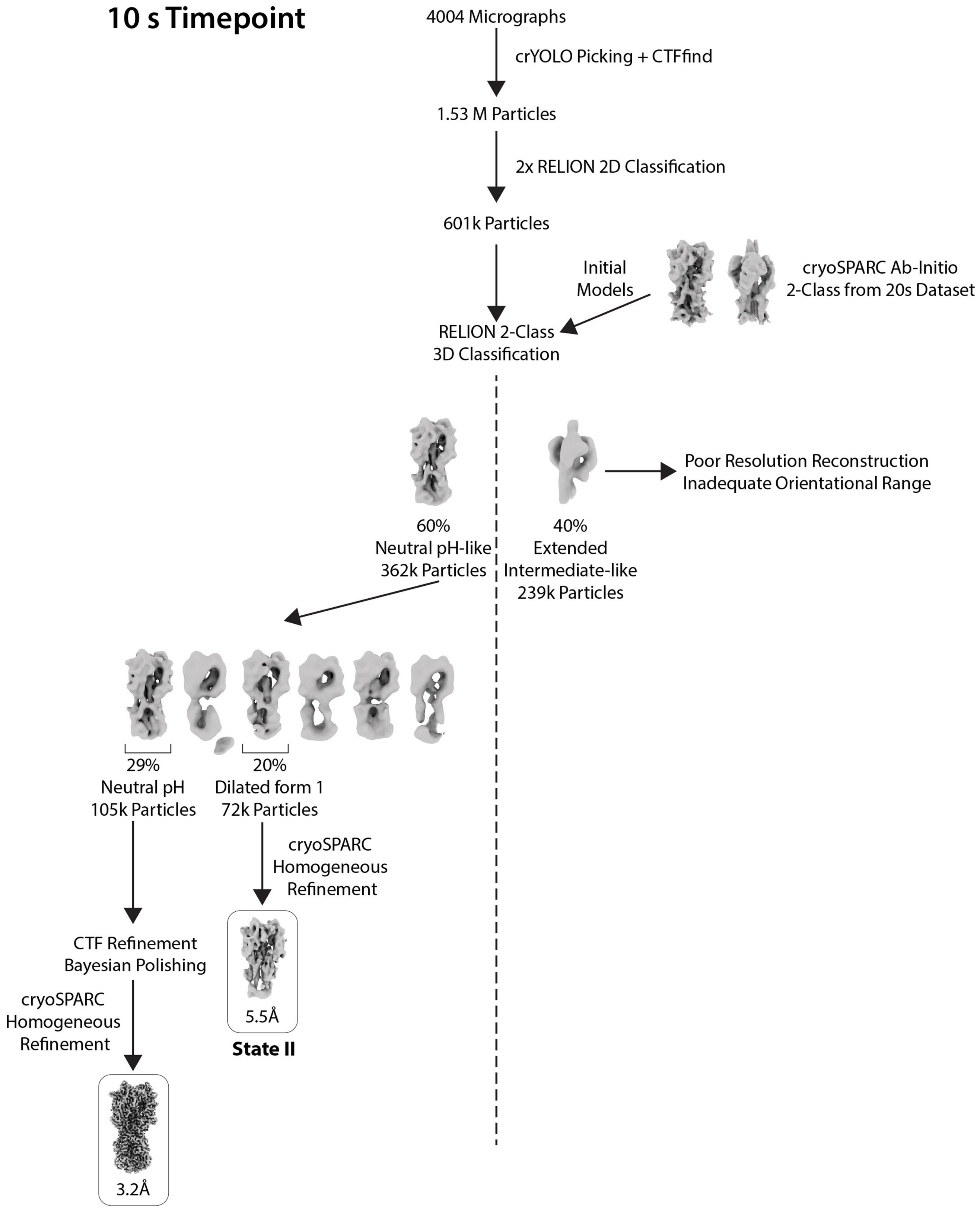
Extended Data Fig. 3 | Flexibility of the membrane-distal domains. The membrane-distal regions of HA1 (roughly residues 45–310) can adopt a range of different orientations compared with the C3 symmetrized structure (state IV). The remainder of HA1 and HA2 adopt a structure with less evidence of flexibility. We examined the flexibility in ten maps generated by asymmetric classification. **a**, The generated models are shown aligned to the symmetrized version of the protein (yellow). **b**, Examples of the most-populated classes are shown on a molecular surface, with HA1 in blue and HA2 in red to emphasize the different locations of these domains when compared with the symmetrized version. **c**, The displacements of the locations of the centroids of each of these mobile domains were measured to the nearest symmetrized monomer, giving 30 data points. These adopt a range of locations, with the main flexibility being in a rotation angle around the threefold axis, with little lateral movement towards and away from the threefold axis. The angles of displacement vary from −15° to +25°.



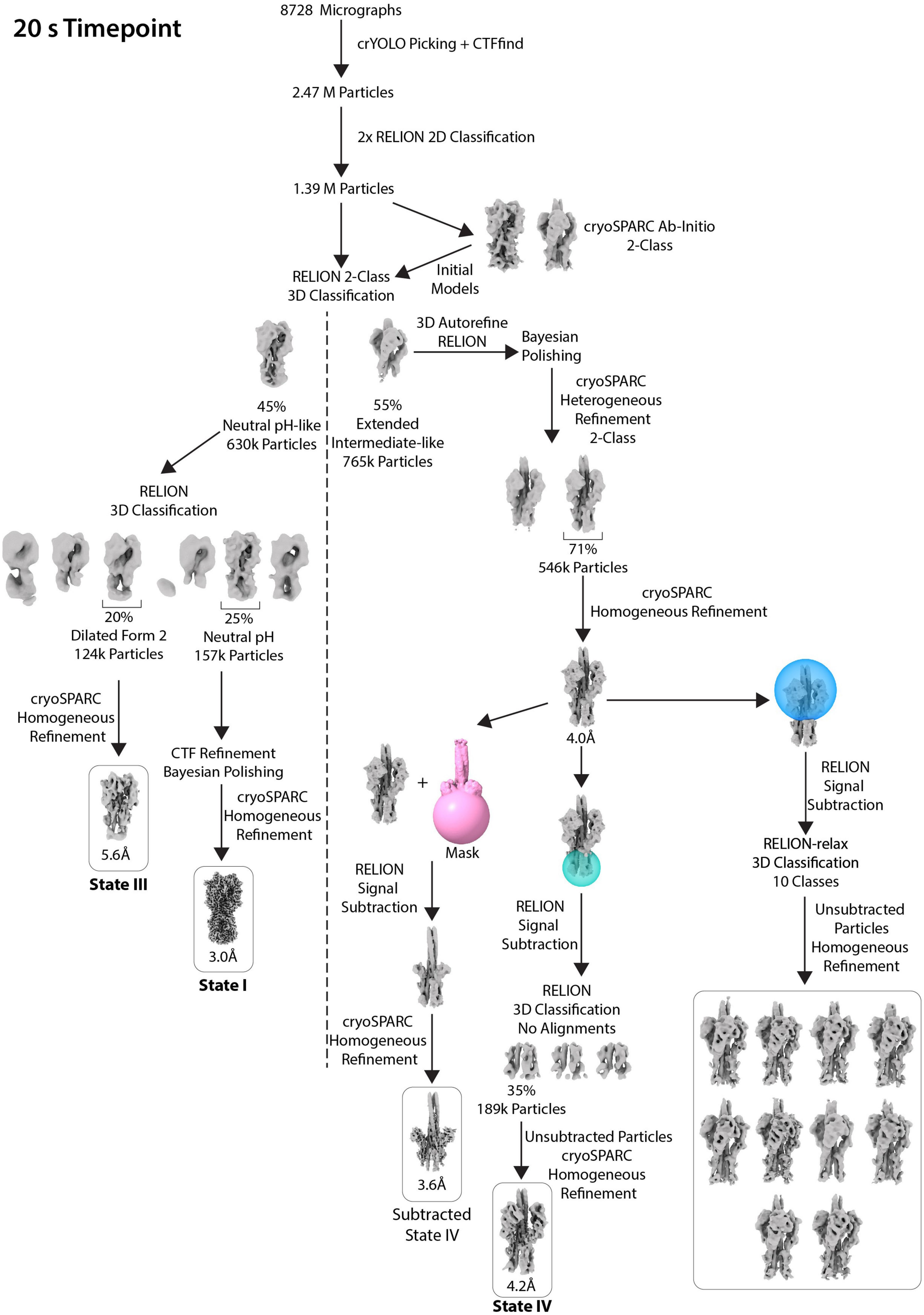
Extended Data Fig. 4 | The 30-loop. a, Potential interactions in the 30-loop are similar in both the neutral-pH (state I) and the extended HA2 (state IV) conformations. There are, however, several changes in the side chains involved in this interaction, owing to the relocation of the short helix of HA2 in the neutral-pH structure. This helix relocation removes an HA2 salt bridge between Arg54 of the short helix and Glu97 of the long helix, as well as the interaction of Lys51 with His106. These rearrangements permit new potential interactions between Thr30 of HA1 with Gln105 of HA2 and His106 of the adjacent HA2 chain. A salt bridge also forms between Lys27 of HA1 and Glu97 of HA2. **b**, Density of the 30-loop and interacting regions of the long helices of HA2 in the extended subtracted structure. **c**, The location and architecture of the 30-loop is conserved in all HAs, including influenza C HEF[37]. In HA, a cluster of conserved hydrophobic-loop residues at positions 26, 34 and 36 packs against the strictly conserved Ile108 in the long α-helix; the strictly conserved Asn104 forms hydrogen bonds with loop residue Lys27 and Lys315 of HA1; and HA2 residue 105, conserved as Gln or Glu, interacts with Thr30. Amino-acid substitutions in the loop (Thr30 to Ser), the short α-helix (Arg54 to Lys) and the long α-helix (Gln105 to Lys and His106 to Ala) that interact with loop residues 28, 29 and 30, respectively, destabilize the mutant HAs, as shown by their elevated fusion pH[17,20]. These observations indicate the functional importance of the loop and suggest its involvement in membrane fusion. Formation of the 180° turn in the extended helix requires removal of the 30-loop from its interactions with Gln105 and His106. The observed unfolding of the loop and its acquisition of susceptibility to protease digestion in stage V are consistent with this requirement and with the suggested role of the loop in supporting the extended helix in states II, III and IV.



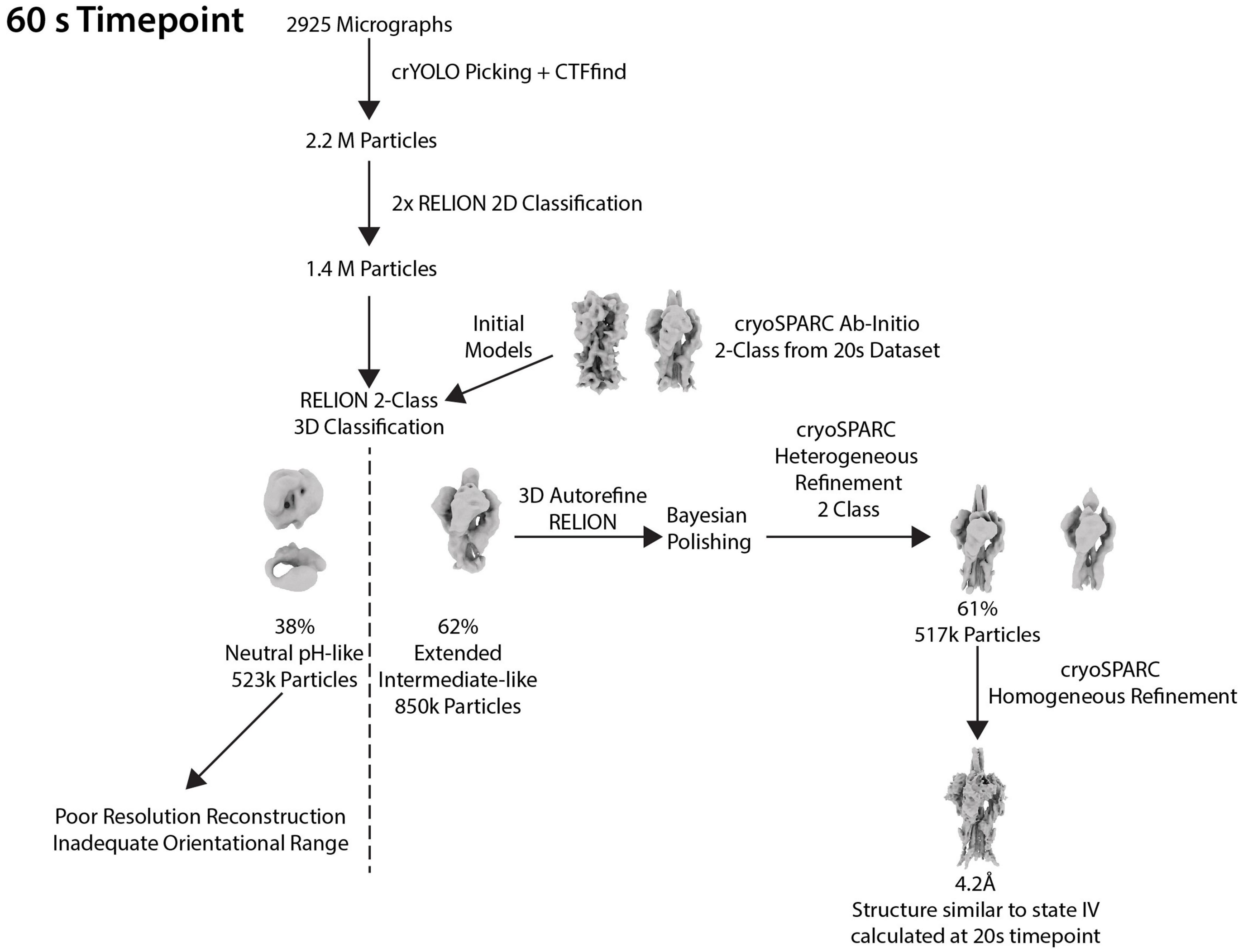
Extended Data Fig. 5 | 2D class averages and representative micrographs for the 10 s, 20 s and 60 s time points. a, 2D classes fall into two broad groups: neutral-pH-like and extended-intermediate-like. Representative classes are shown for each group for each time point, with the percentages indicating the overall numbers of particles that fall into these two broad groups. There is a general trend towards increasing numbers of extended-intermediate-like classes as the time point increases, with a decrease in numbers of neutral-pH-like classes. There is limited diversity in the orientations seen for the 10 s extended-intermediate-like and the 60 s neutral-pH-like classes. **b**, Representative micrographs (scale bars, 50 nm). Numbers of micrographs collected: 10 s, *n* = 4,004; 20 s, *n* = 8,728; 60 s, *n* = 2,925. **c**, SDS–PAGE of the HA sample used for experiments.



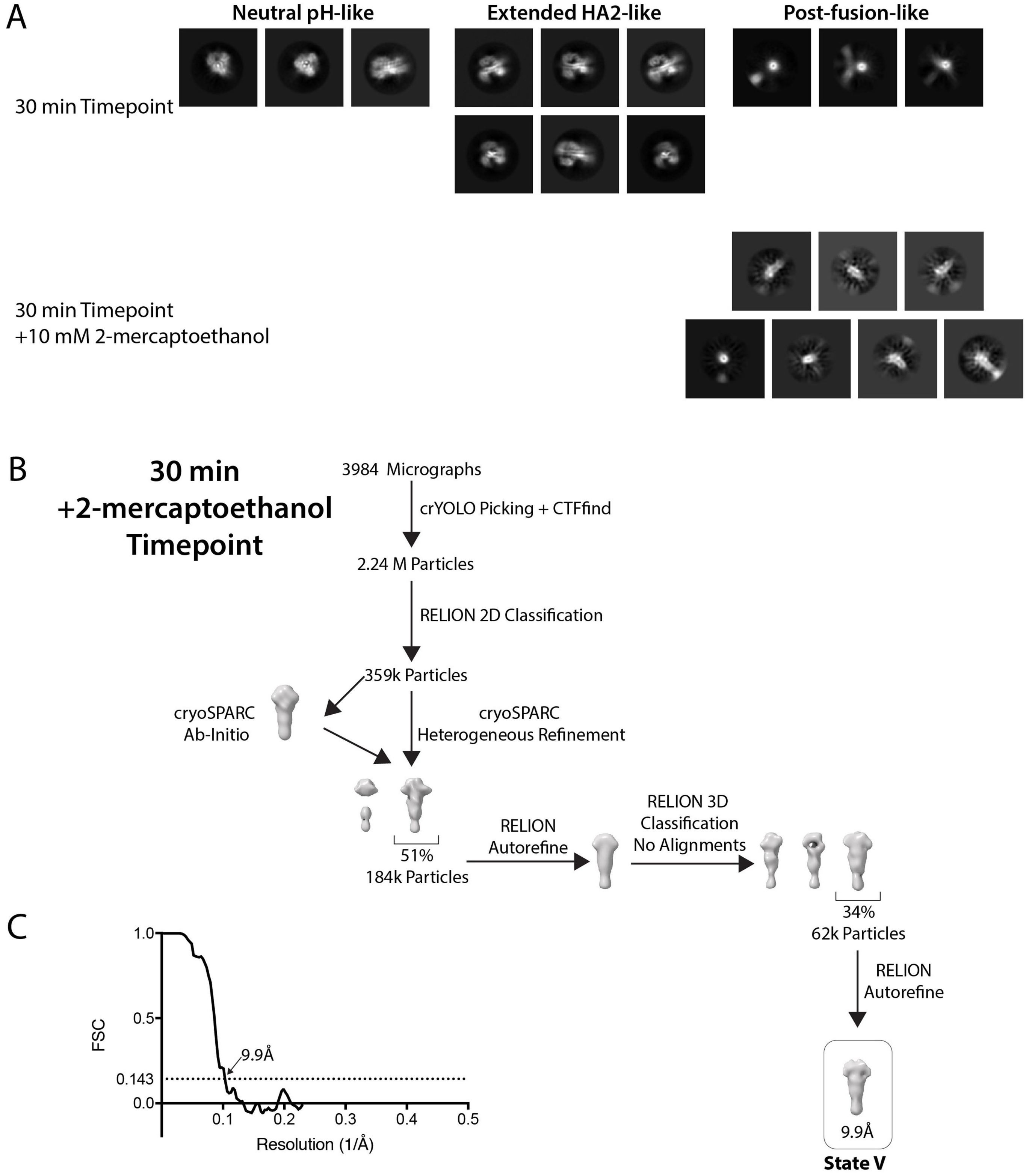
Extended Data Fig. 6 | Classification scheme for the 10 s time point. The figure shows the workflow for image processing to determine structures at 10 s.



Extended Data Fig. 7 | Classification scheme for the 20 s time point. The figure shows the workflow for image processing to determine structures at 20 s.



Extended Data Fig. 8 | Classification scheme for the 60 s time point. The figure shows the workflow for image processing to determine structures at 60 s.



Extended Data Fig. 9 | Data processing for 30 min time points. a, 2D classes from the 30 min time point show the presence of the extended HA2 (state IV) and the post-fusion conformations, as well as a minor population of neutral-pH conformations. The post-fusion averages are interpretable only for the end-on view down the threefold axis, where this conformation forms a characteristic dense ring-like shape. This limited range of interpretable views is probably due to the, now flexibly linked, HA1 domains increasing the structural heterogeneity. The interpretability of this species could be improved by the addition of 10 mM of 2-mercaptoethanol to detach the disulfide-linked HA1 domains. **b**, Classification scheme for the 30-min time point supplemented with 2-mercaptoethanol. **c**, FSC curve for the final refined structure, with resolution criterion FSC = 0.143 shown.

Extended Data Table 1 | Cryo-EM statistics for map and model refinement

	Pre-Fusion pH8 (EMDB-10696) (PDB 6Y5G)	Pre-Fusion pH5 (EMDB-10697) (PDB 6Y5H)	Dilated Form 1 (EMDB-10698) (PDB 6Y5I)	Dilated Form 2 (EMDB-10699) (PDB 6Y5J)	Full Extended HA2 (EMDB-10700) (PDB 6Y5K)	Subtracted Extended HA2 (EMDB-10701) (PDB 6Y5L)	Post Fusion (EMDB-10702)
State Number	-	I	II	III	IV	IV	V
Timepoint	-	20 s	10 s	20 s	20 s	20 s	30 min + 2-ME
Data collection and processing							
Magnification	75 000	75 000	75 000	75 000	75 000	75 000	75 000
Voltage (kV)	300	300	300	300	300	300	300
Electron exposure (e–/Å^2)	33.9	33.9	33.9	33.9	33.9	33.9	33.9
Defocus range (μm)	1.5-3	1.5-3	1.5-3	1.5-3	1.5-3	1.5-3	1.5-3
Pixel size (Å)	1.09	1.09	1.09	1.09	1.09	1.09	1.09
Symmetry imposed	C3	C3	C1	C1	C3	C3	C3
Initial particle images (no.)	643 k	1.39 M	1.53 M	1.39 M	1.39 M	1.39 M	2.24 M
Final particle images (no.)	200 k	157 k	72 k	124 k	189 k	546 k	62 k
Map resolution (Å) FSC threshold 0.143	3.0	3.0	5.5	5.6	4.2	3.6	9.9
Refinement Software	CryoSPARC	CryoSPARC	CryoSPARC	CryoSPARC	CryoSPARC	CryoSPARC	RELION
Map resolution range (Å)	2.6-3.4	2.6-3.4	5-9	5-9	3-5	3-7	-
Refinement							
Initial model used (PDB code)	2YPG	2YPG	2YPG	2YPG	2YPG & 1QU1	2YPG & 1QU1	-
Map sharpening *B* factor (Å^2)	-172	-170	-283	-322	-170	-225	-
Model composition							
Non-hydrogen atoms	12036	11994	11928	9609	10617	3855	-
Protein residues	1470	1470	1470	1185	1335	477	-
R.m.s. deviations							
Bond lengths (Å)	0.005	0.004	0.004	0.005	0.005	0.004	-
Bond angles (°)	0.553	0.959	0.987	1.000	1.049	0.999	-
Validation							
MolProbity score	1.48	1.35	1.54	1.82	1.65	1.78	-
Clashscore	3.69	2.81	4.31	10.88	10.23	5.46	-
Poor rotamers (%)	0.94	0.70	0.47	0.94	0	0	-
Ramachandran plot							
Favored (%)	95.4	95.9	95.3	96.2	97.35	92.16	-
Allowed (%)	4.6	4.1	4.7	3.8	2.65	7.84	-
Disallowed (%)	0	0	0	0	0	0	-

nature research

Corresponding author(s): Donald J. Benton, Peter B. Rosenthal

Last updated by author(s): Feb 25, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a | Confirmed

☐ ☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement

☒ ☐ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly

☒ ☐ The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.

☒ ☐ A description of all covariates tested

☒ ☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons

☒ ☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)

☒ ☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted
Give P values as exact values whenever suitable.

☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings

☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes

☒ ☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	CryoEM data collected using Thermo Scientific EPU v1.9.0
Data analysis	CryoEM data processed using following packages: RELION-3.0, Relion-2.0-relax, cryoSPARC v2.11, CTFfind4 v.4.1.10, MotionCor2 v.1.2.6, crYOLO v1.4, Coot v.0.9-pre, PHENIX v.1.17, UCSF Chimera v.1.12, UCSF ChimeraX v.0.5

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Maps and models have been deposited in the Electron Microscopy Data Bank, http://www.ebi.ac.uk/pdbe/emdb/ (Accession Nos. EMD-10696, EMD-10697, EMD-10698, EMD-10699, EMD-10700, EMD-10701, EMD-10702). Models have been deposited in the Protein Data Bank, https://www.ebi.ac.uk/pdbe/ (PDB ID codes: 6Y5G, 6Y5H, 6Y5I, 6Y5J, 6Y5K and 6Y5L).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	All datasets consist of several thousands of independent images, as described in Methods. The number of images were sufficient to achieve the reported resolution, according to the most commonly reported resolution measures in cryoEM.
Data exclusions	cryoEM single particles were included and excluded using standard image processing classification techniques. Details of numbers of selected images are shown in Extended Data Table 1.
Replication	Structures were determined using half datasets, according to standard procedures in cryoEM. Preliminary images of at least 3 samples for each condition were consistent in low resolution visual assessment. A single grid was selected for high-resolution data collection and analysis.
Randomization	Not applicable to this study, as samples were not assigned to experimental groups and data were collected and processed according to standard techniques for cryoEM.
Blinding	Not applicable to this study, as there was no experimental group allocation in data collection and analysis.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Transformation of naked mole-rat cells

https://doi.org/10.1038/s41586-020-2410-x

Received: 11 April 2018

Accepted: 22 April 2020

Fazal Hadi[1,2], **Yavuz Kulaberoglu**[1], **Kyren A. Lazarus**[1,2], **Karsten Bach**[1,2], **Rosemary Ugur**[1,2], **Paul Beattie**[1], **Ewan St John Smith**[1]✉ & **Walid T. Khaled**[1,2,3]✉

ARISING FROM Tian, X. et al. *Nature* https://doi.org/10.1038/nature12234 (2013)

The naked mole rat (NMR), *Heterocephalus glaber*, is a mouse-sized subterranean rodent that is native to East Africa and is used in research for the potential development of therapeutics because of its unusual physiology[1,2], long lifespan[3] and cancer resistance[2,4]. In a previous study, Tian et al.[5] reported that the cancer resistance of NMRs is mediated by high-molecular-mass hyaluronan produced by NMR cells and showed that wild-type NMR cells, but not cells in which hyaluronan expression is perturbed, are resistant to transformation by SV40 large T antigen (encoded by *SV40LT*) and oncogenic HRAS ($HRAS^{G12V}$)—a combination of oncogenes that is sufficient to transform mouse and rat fibroblasts[6,7]. Here we developed a number of lentiviral vectors to deliver both of these oncogenes and generated 106 different cell lines from 5 different tissues and 11 different NMRs and show that, in contrast to the previous study[5], NMR cells are susceptible to oncogenic transformation by *SV40LT* and $HRAS^{G12V}$. Our data thus suggest that a non-cell autonomous mechanism underlies the remarkable cancer resistance of NMRs and that identifying this non-cell autonomous mechanism could have important implications for our understanding of cancer development in humans.

Few studies have attempted to explain the mechanisms that underlie the cancer resistance of NMRs[8–11]. Previously, Tian et al.[5] reported that the expression of high-molecular-mass hyaluronan in NMR cells mediates its cancer resistance. With the publication of the NMR genome[12,13] and advances in the gene editing through the use of CRISPR–Cas9 technology[14], we set out to comprehensively identify genes in the NMR that are responsible for its cancer resistance. We based our approach on the observation by Tian et al.[5] that NMR cells are resistant to transformation by *SV40LT* and $HRAS^{G12V}$, a combination of oncogenes that is sufficient to transform mouse and rat fibroblasts[6,7].

We therefore generated a number of lentiviral vectors that would enable us to deliver single-guide RNAs (sgRNAs) together with *SV40LT* and/or $HRAS^{G12V}$ under the control of two different promoters (*Pgk1* and *EF1A*, Extended Data Fig. 1a). On the basis of the study by Tian et al.[5], we expected that NMR cells transduced with any of our vectors would fail to grow in anchorage-independent conditions unless a further gene was perturbed, thus providing us with a defined system for our studies. Using our lentiviral vectors, we generated cell lines from 5 different tissues (intestine, kidney, pancreas, lung and skin) from 11 different NMRs (Extended Data Fig. 1b). We then tested 79 different cell lines for anchorage-independent growth using the protocol described by Tian et al.[5]. As expected, and in line with previous reports[5,8], primary cell lines and those transduced with *SV40LT* alone failed to grow in soft agar, even after six weeks in culture. By contrast, NMR cell lines expressing both *SV40LT* and $HRAS^{G12V}$ formed robust colonies in soft agar (Fig. 1a and Extended Data Fig. 1c–f). These results were reproducible for all cell lines tested, irrespective of the animal or promoter used to drive the expression of the exogenous *SV40LT* and $HRAS^{G12V}$ genes.

On the basis of these results, we tested the tumorigenic potential of the cell lines expressing *SV40LT* and $HRAS^{G12V}$. Untransduced parental cell lines and their respective, *SV40LT*-expressing (hereafter, *SV40LT*) and *SV40LT*- and $HRAS^{G12V}$-overexpressing cell lines (hereafter, *SV40LT;*$HRAS^{G12V}$) were injected subcutaneously into non-obese diabetic/severe combined immunodeficiency/IL2Rγ (NSG) immunocompromised mice and monitored the mice for tumour growth. As early as day 5 after the injection of NMR cells, tumour masses were detected in mice bearing NMR cell lines that expressed *SV40LT;*$HRAS^{G12V}$ (Fig. 1b and Extended Data Fig. 2a, b). By day 40 after cell injection, all mice bearing transformed NMR cell lines had been euthanized and the tumours collected, whereas mice injected with control (primary or *SV40LT* alone) NMR cells did not show any signs of tumour growth, even at 60 days after injection when the experiment was terminated (Fig. 1b and Extended Data Fig. 2a, b).

In an attempt to explain the discrepancy between our findings and those of Tian et al.[5], the authors provided us with a detailed list of materials and protocols that we used throughout our study. In addition, the authors also provided us with their primary and *SV40LT;*$HRAS^{G12V}$-transfected cell lines to test in our laboratory. Similar to our primary cells from multiple tissues, primary skin cell lines obtained from Tian et al.[5] were transformed by our *SV40LT*- and $HRAS^{G12V}$-expressing vectors and formed robust colonies in vitro and tumours in mouse xenograft assays (Fig. 1c, d and Extended Data Fig. 2c, d); thus, excluding the possibility that variation in NMR colonies caused the observed differences. Notably, the *SV40LT;*$HRAS^{G12V}$-transfected cell line from Tian et al.[5] also formed colonies in vitro and tumours in xenograft assays, which we have quantified (Fig. 1d, and Extended Data Fig. 2d). However, these cells had a lower clonogenic efficiency and longer tumour latency.

To further investigate these contradictory results, we generated lentiviral vectors in which *SV40LT;*$HRAS^{G12V}$ expression was driven by the same promoters used by Tian et al.[5] (that is, *SV40E* and *CMVIE*; Extended Data Fig. 1a). We also decided to use exactly the same basal media as Tian et al.[5] (EMEM) to culture the cells, thus excluding any potential metabolic effects mediated by DMEM. The soft-agar and xenograft assays again confirmed our original observation that NMR cells can be transformed by expression of *SV40LT;*$HRAS^{G12V}$ and that the use of different promoters or basal medium had no role in the transformation of NMR cells (Figs. 1e, f, 2a and Extended Data Fig. 2e). In further experiments, we also titrated the amount of virus used to ensure a single transduction event per cell. In addition, we transfected cells with a linearized vector to completely avoid the use of viruses. Neither of these approaches changed the outcome of the soft-agar and xenograft assays, thus demonstrating that the source of the discrepancy between our results and those of Tian et al.[5] was not due to lentivirus-mediated mutagenesis (Fig. 2b, c and Extended Data Fig. 2e, f).

---

[1]Department of Pharmacology, University of Cambridge, Cambridge, UK. [2]Cancer Research UK Cambridge Cancer Centre, Cambridge, UK. [3]Cambridge Stem Cell Institute, Cambridge, UK. ✉e-mail: es336@cam.ac.uk; wtk22@cam.ac.uk

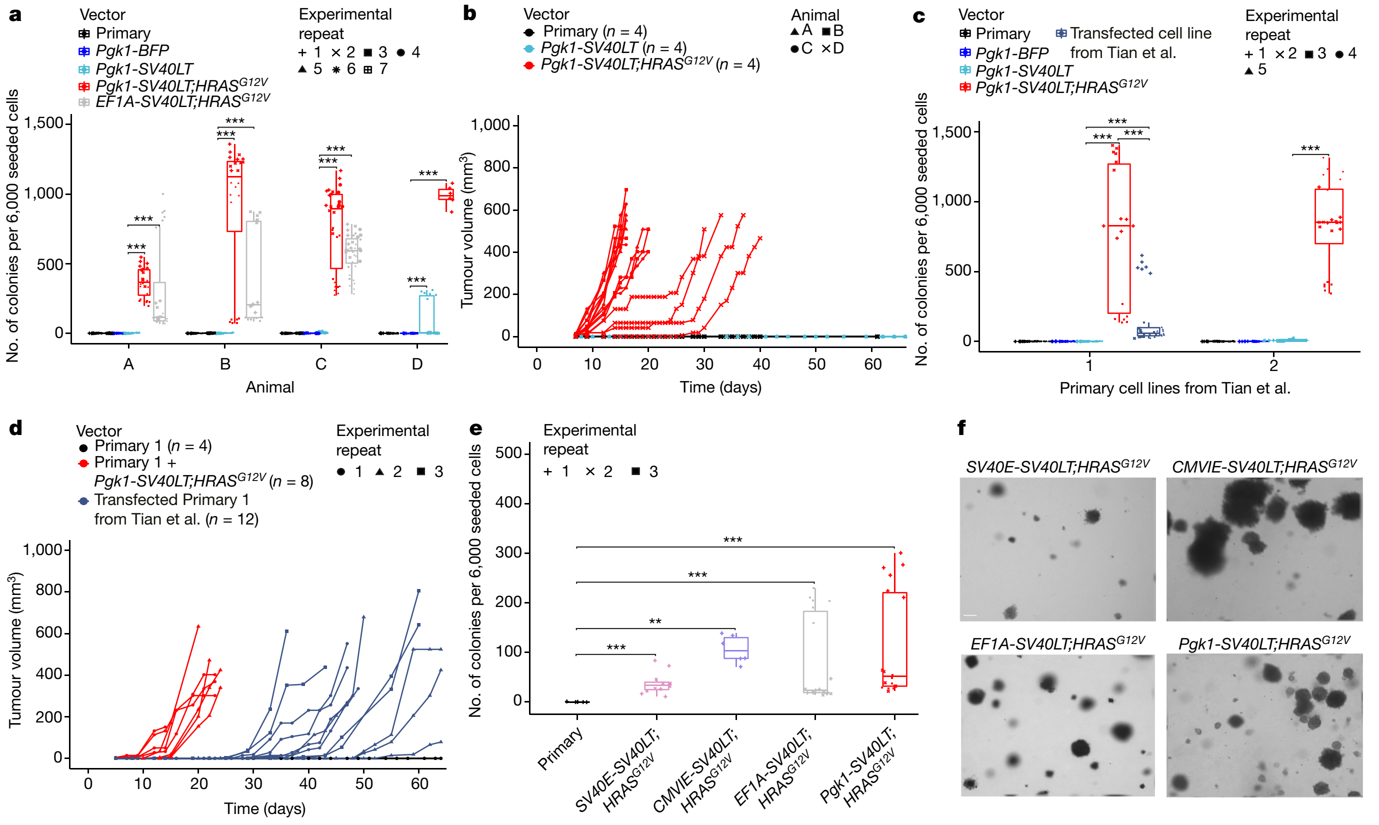


Fig. 1 | NMR cells can be transformed by *SV40LT* and *HRAS^{G12V}*. **a**, Quantification of soft-agar colonies from skin cell lines generated from four different NMRs. Different shapes represent different experimental repeats. A single experiment was repeated up to seven times and each experimental repeat included six technical replicates. Each data point represents the number of colonies observed from an individual technical replicate. In total, more than 7,000 fields of view (non-overlapping images) were analysed. BFP, blue-fluorescent protein. **b**, Quantification of xenograft tumour growth in NSG mice injected with primary cells or cells transduced with *Pgk1-SV40LT* or *Pgk1-SV40LT;HRAS^{G12V}*. Each cell line was injected into four mice; each mouse is represented by a single line. Cell lines were derived from four different NMRs (represented by different shapes). **c**, Quantification of soft-agar colonies grown using NMR cells from Tian et al.[5]. Colours represent cells expressing different vectors introduced using lentiviral particles or transfection. Different shapes represent individual experimental repeats. A single experiment was repeated up to five times and each experimental repeat included six technical replicates. Each data point represents the number of colonies observed from an individual technical replicate. In total, more than 2,500 fields of view (non-overlapping images) were analysed. **d**, Quantification of xenograft tumour growth in NSG mice injected with cells from Tian et al.[5]. Colours represent different vectors introduced using lentiviral particles or transfections. Shapes represent different experimental repeats. In every repeat, each cell line was injected into four mice; each mouse is represented by a single line. **e**, Quantification of soft-agar colonies from primary NMR skin cells or their respective counterpart cell lines generated by introducing *SV40LT;HRAS^{G12V}* (under different promoters) using titrated lentiviral particles. Colours represent different vectors and shapes represent different experimental repeats. Each experiment was repeated up to three times and each experimental repeat included six technical replicates. Each data point represents the number of colonies observed from an individual technical replicate. In total, more than 7,500 fields of view (non-overlapping images) were analysed. **a**, **c**, **e**, Data were analysed using Wilcoxon rank-sum tests; $^{**}P \leq 0.01$; $^{***}P \leq 0.0001$. Box plots are shown as follows: centre line, median; box limits, upper and lower quartiles; whiskers, 1.5× the interquartile range. **f**, Representative images of colonies observed in the soft-agar assay (the quantification is shown in **e**) from cell lines generated by introducing *SV40LT;HRAS^{G12V}* (under different promoters). Scale bar, 100 µm.

After excluding different promoters, media and lentivirus-mediated mutagenesis as sources of discrepancy, we set out to repeat the transfection with the same transient expression vectors reported by Tian et al.[5]. We used exactly the same protocol as reported in their paper[5] to transfect primary NMR and mouse skin cells. The cells were then seeded 24 h later in soft agar and after 6 weeks, we were unable to detect robust colonies from the NMR cells (Fig. 2d). However, in contrast to the previously reported results[5], we were also unable to detect robust colonies from the mouse cells (Fig. 2e). This can be explained by the fact that Tian et al.[5]. used multiple transient expression vectors without any antibiotic selection to introduce *SV40LT* and *HRAS^{G12V}* vectors into primary cells. This is an extremely inefficient method for introducing exogenous oncogenes given that the transformation assay lasts 6 weeks.

Finally, we performed RNA sequencing of NMR and mouse cell lines to explore the transcriptional response of these cell lines to the introduction of *SV40LT* and *SV40LT;HRAS^{G12V}*. Principal component analysis (PCA) of a set of approximately 16,000 orthologues[15] showed a grouping by species along the first principal component (PC1). Notably, samples appear to be ordered based on the level of transformation along second principal component (PC2) from untransformed to *SV40LT* to *SV40LT;HRAS^{G12V}*, independent of the species (Fig. 2f). Further examination showed that genes with positive loadings for PC2 were significantly enriched for cell adhesion and genes with negative loadings were involved in proliferation (Extended Data Fig. 3a). There were few transcriptional differences between our transformed cell line and the transfected cell line of Tian et al.[5] (227 differentially expressed genes, false-discovery rate (FDR)-adjusted $P = 0.01$) compared with the effect of transformation (1,738 differentially expressed genes, FDR-adjusted $P = 0.01$; transformed compared with untransformed), which is also illustrated by the grouping in the PCA (Fig. 2f and Extended

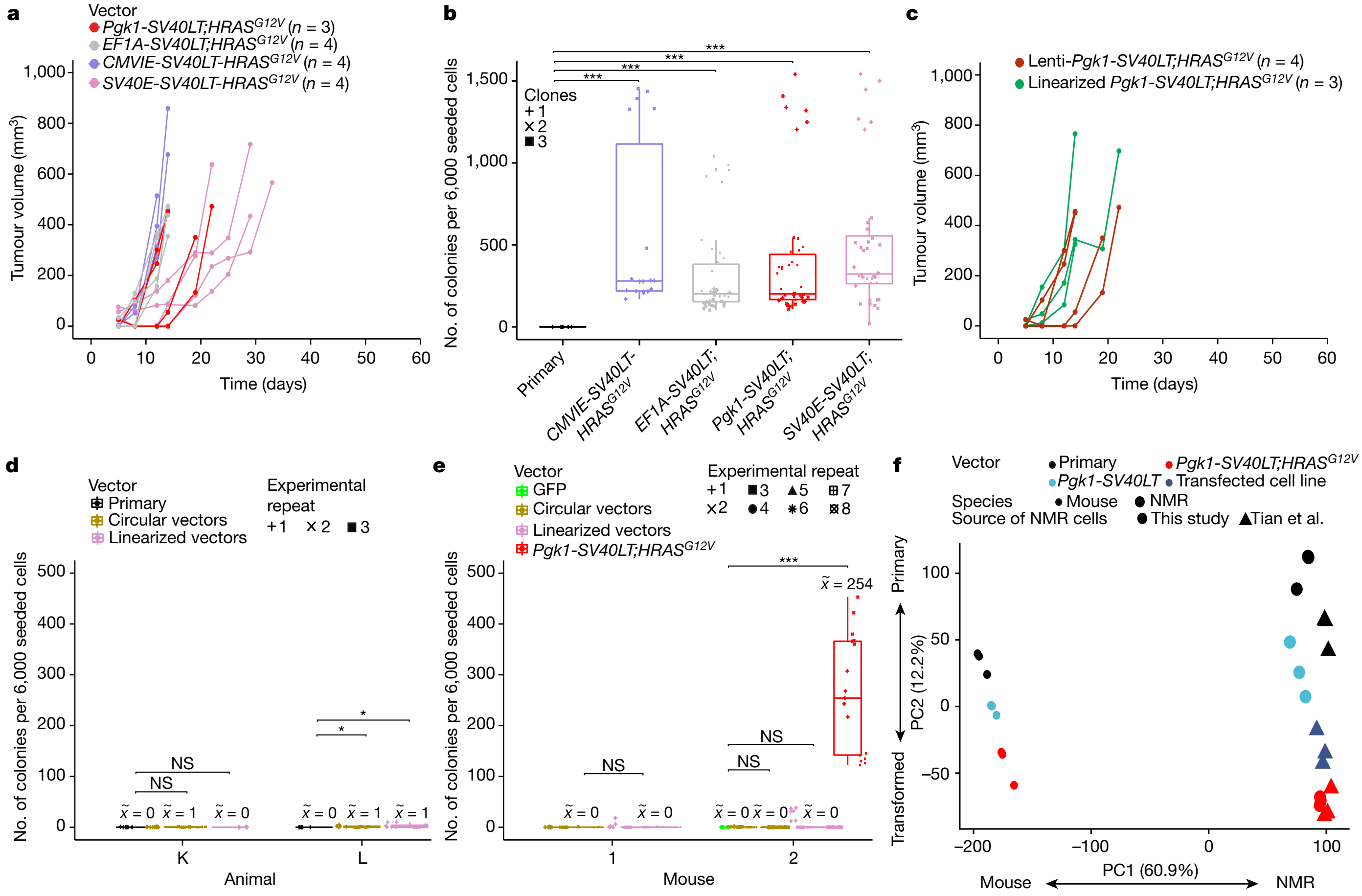


Fig. 2 | NMR transformation is not dependent on the promoter, vector or culturing medium used. a, Quantification of xenograft tumour growth in NSG mice injected with NMR skin cell lines generated by introducing *SV40LT;HRAS*^{G12V} (under different promoters; *SV40E* and *CMVIE* were also used by Tian et al.[5]; or *Pgk1* or *EF1A*) via lentiviral particles. Each cell line was injected into four mice (or three in the case of *Pgk1-SV40LT;HRAS*^{G12V}); each mouse is represented by a single line. Colours represent different vectors. **b**, Quantification of soft-agar colonies from primary NMR skin cells or their respective counterpart cell lines generated by introducing *SV40LT;HRAS*^{G12V} (under different promoters: *CMVIE*, *SV40E*, *EF1A* or *Pgk1*) through the transfection of linearized vectors. For each cell line, up to three different independent clones were generated (represented by different shapes). Each experiment was repeated up to three times and each experimental repeat included six technical replicates. Each data point represents the number of colonies observed in a single technical replicate. In total, more than 21,000 fields of view (non-overlapping images) were analysed. **c**, Quantification of xenograft tumour growth in NSG mice injected with NMR skin cell lines generated by introducing *Pgk1-SV40LT;HRAS*^{G12V} using lentiviral particles or transfection of linearized plasmids (represented by different colours). Each cell line was injected into at least three mice; each mouse is represented by a single line. **d**, Quantification of soft-agar colonies from primary NMR skin cells or primary NMR skin cells transfected with pmaxEGFP and circular or linearized pCMV-RasV12 and pSG5-largeT (plasmids used by Tian et al.[5]). NMR skin cells were derived from two different animals; each experiment was repeated up to three times and each experimental repeat included six technical replicates (represented by different shapes). Each data point indicates the number of colonies observed in a single technical replicate. In total, 7,800 fields of view (non-overlapping images) were analysed. **e**, Quantification of soft-agar colonies from mouse skin cells transfected with pmaxEGFP and circular or linearized pCMV-RasV12 and pSG5-largeT (plasmids used by Tian et al.[5]) or transduced with *Pgk1-SV40LT;HRAS*^{G12V} (represented by different colours). Mouse skin cells were derived from two animals; each experiment was repeated up to eight times and each experimental repeat included six independent technical replicates (represented by different shapes). Each data point indicates the number of colonies observed in a single technical replicate. In total, more than 25,000 fields of view (non-overlapping images) were analysed. **d**, **e**, Circular vectors, a mixture of circular pCMV-RasV12, pSG5-largeT and pmaxEGFP plasmids; linearized vectors, a mixture of linearized pCMV-RasV12 and pSG5-largeT plasmids plus circular pmaxEGFP plasmid. **f**, PCA of mouse and NMR primary skin cells or their respective cell lines, which express *SV40LT* or *SV40LT;HRAS*^{G12V}, are represented by different colours. Species are separated along PC1, whereas cells based on transformation status are separated along PC2; primary cells have a high PC2 value compared with *Pgk1-SV40LT;HRAS*^{G12V}-transformed cells. Note the proximity of the transfected NMR skin cell line from Tian et al.[5] to our *Pgk1-SV40LT;HRAS*^{G12V}-transformed cells. **b**, **d**, **e**, Data were analysed using Wilcoxon rank-sum tests; $^{*}P \leq 0.05$; $^{***}P \leq 0.0001$; NS, not significant. Box plots are shown as follows: centre line, median; box limits, upper and lower quartiles; whiskers, 1.5× the interquartile range. **d**, **e**, $\tilde{x}$, median number of colonies.

Data Fig. 3b). Together, these data suggest that cells from both NMRs and mice respond to the introduction of *SV40LT;HRAS*^{G12V} by downregulating processes that are involved in cell adhesion and by upregulating proliferation, which is in accordance with the observed phenotype of anchorage-independent growth.

On the basis of our results, we conclude that NMR cells are not resistant to transformation by *SV40LT* and *HRAS*^{G12V}. Our data therefore suggest that the key mechanisms that underlie the cancer resistance of NMRs are non-cell autonomous and instead might be explained by a unique microenvironment and/or immune system. This is supported by a recent report in which the unique immune system of NMRs is described[15]. It is worth noting that we have also found that NMR hyaluronan has unique physical characteristics[16]. However, in light of our results it remains to be seen whether these physical characteristics of hyaluronan have a role in cancer resistance. Therefore, it is important that the field is aware of our findings so that efforts in understanding the remarkable biology and cancer resistance of NMRs are guided in the right direction.

Methods

A full list of materials, methods and experimental protocols can be found in the Supplementary Methods.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The RNA-sequencing data are available from the European Nucleotide Archive (accession number E-MTAB-8932). Raw western blot data are provided as Supplementary Fig. 1. The authors declare that all remaining supporting data are available within the paper and the Supplementary Information or from the corresponding authors upon reasonable request.

Code availability

The source code used for analysis of the RNA-sequencing data is available at https://github.com/KaBach/NMR.

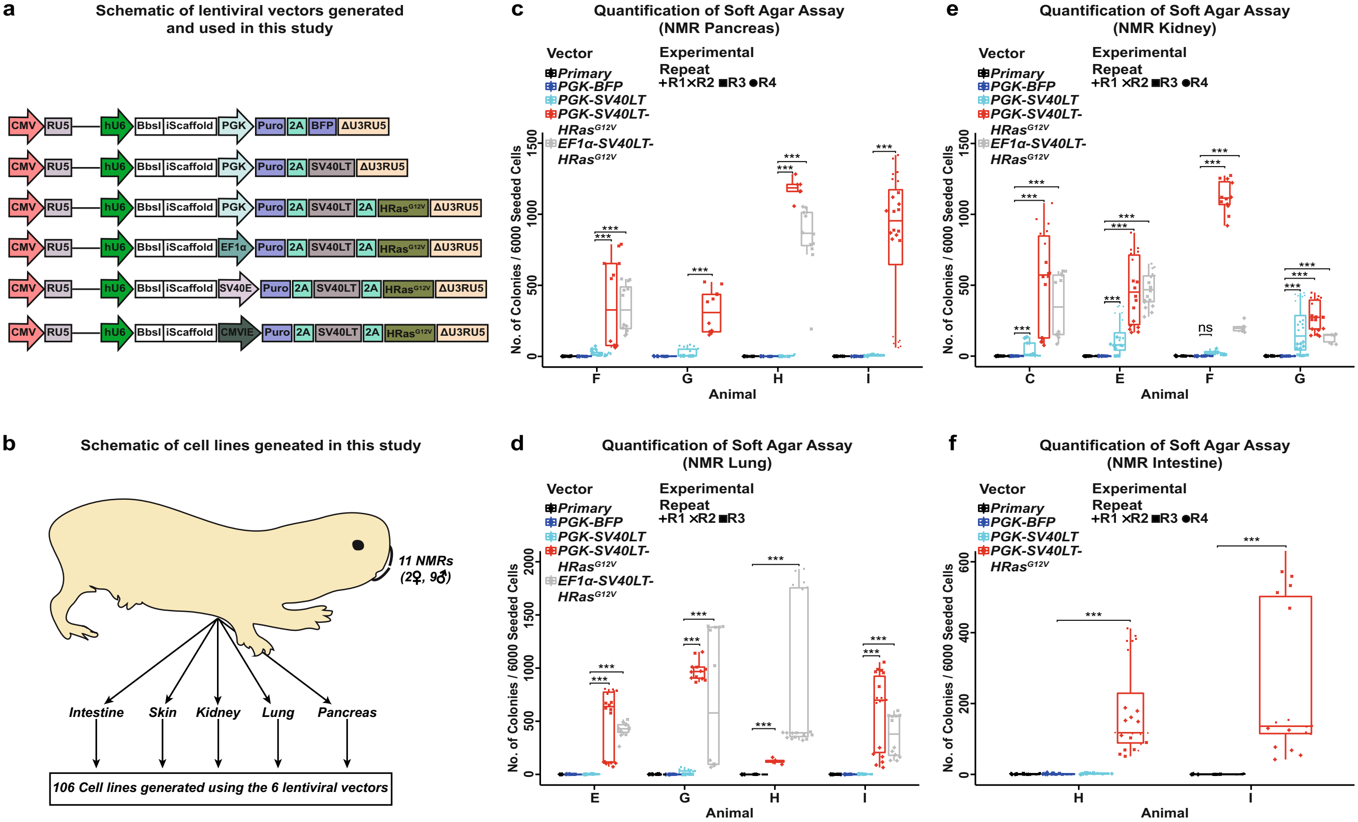
1. Park, T. J. et al. Fructose-driven glycolysis supports anoxia resistance in the naked mole-rat. *Science* **356**, 307–311 (2017).
2. Schuhmacher, L.-N., Husson, Z. & Smith, E. S. J. The naked mole-rat as an animal model in biomedical research: current perspectives. *Open Access Anim. Physiol*. **7**, 137–148 (2015).
3. Ruby, J. G., Smith, M. & Buffenstein, R. Naked mole-rat mortality rates defy Gompertzian laws by not increasing with age. *eLife* **7**, e31157 (2018).
4. Buffenstein, R. Negligible senescence in the longest living rodent, the naked mole-rat: insights from a successfully aging species. *J. Comp. Physiol. B* **178**, 439–445 (2008).
5. Tian, X. et al. High-molecular-mass hyaluronan mediates the cancer resistance of the naked mole rat. *Nature* **499**, 346–349 (2013).
6. Rangarajan, A., Hong, S. J., Gifford, A. & Weinberg, R. A. Species- and cell type-specific requirements for cellular transformation. *Cancer Cell* **6**, 171–183 (2004).
7. Michalovitz, D., Fischer-Fantuzzi, L., Vesco, C., Pipas, J. M. & Oren, M. Activated Ha-ras can cooperate with defective simian virus 40 in the transformation of nonestablished rat embryo fibroblasts. *J. Virol*. **61**, 2648–2654 (1987).
8. Seluanov, A. et al. Hypersensitivity to contact inhibition provides a clue to cancer resistance of naked mole-rat. *Proc. Natl Acad. Sci. USA* **106**, 19352–19357 (2009).
9. Miyawaki, S. et al. Tumour resistance in induced pluripotent stem cells derived from naked mole-rats. *Nat. Commun*. **7**, 11471 (2016).
10. Tian, X. et al. *INK4* locus of the tumor-resistant rodent, the naked mole rat, expresses a functional p15/p16 hybrid isoform. *Proc. Natl Acad. Sci. USA* **112**, 1053–1058 (2015).
11. Liang, S., Mele, J., Wu, Y., Buffenstein, R. & Hornsby, P. J. Resistance to experimental tumorigenesis in cells of a long-lived mammal, the naked mole-rat (*Heterocephalus glaber*). *Aging Cell* **9**, 626–635 (2010).
12. Kim, E. B. et al. Genome sequencing reveals insights into physiology and longevity of the naked mole rat. *Nature* **479**, 223–227 (2011).
13. Fang, X. et al. Adaptations to a subterranean environment and longevity revealed by the analysis of mole rat genomes. *Cell Rep*. **8**, 1354–1364 (2014).
14. Sander, J. D. & Joung, J. K. CRISPR–Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol*. **32**, 347–355 (2014).
15. Hilton, H. G. et al. Single-cell transcriptomics of the naked mole-rat reveals unexpected features of mammalian immunity. *PLoS Biol*. **17**, e3000528 (2019).
16. Kulaberoglu, Y. et al. The material properties of naked mole-rat hyaluronan. *Sci. Rep*. **9**, 6632 (2019).

Acknowledgements We thank M. Waterhouse for reading of and comments on the manuscript, S. Nik-Zainal for access to EVOS FLA2 and the staff at the Sanger Institute, Research Service Facility (RSF) for their assistance. F.H. is funded by a Gates Cambridge Trust PhD scholarship. Y.K. is funded by a CRUK multidisciplinary award to E.S.J.S. K.A.L. is funded by a CRUK career establishment award and The Isaac Newton Trust Grant to W.T.K. K.B. is funded by CRUK Cambridge Centre PhD studentship. R.U. is funded by NC3Rs PhD studentship. This work was funded by donations from P.B., Magdalene College (Cambridge), The Isaac Newton Trust Grant (16.38c) and a CRUK Grant (C56829/A22053) to E.S.J.S. and a CRUK grant (C47525/A17348) to W.T.K.

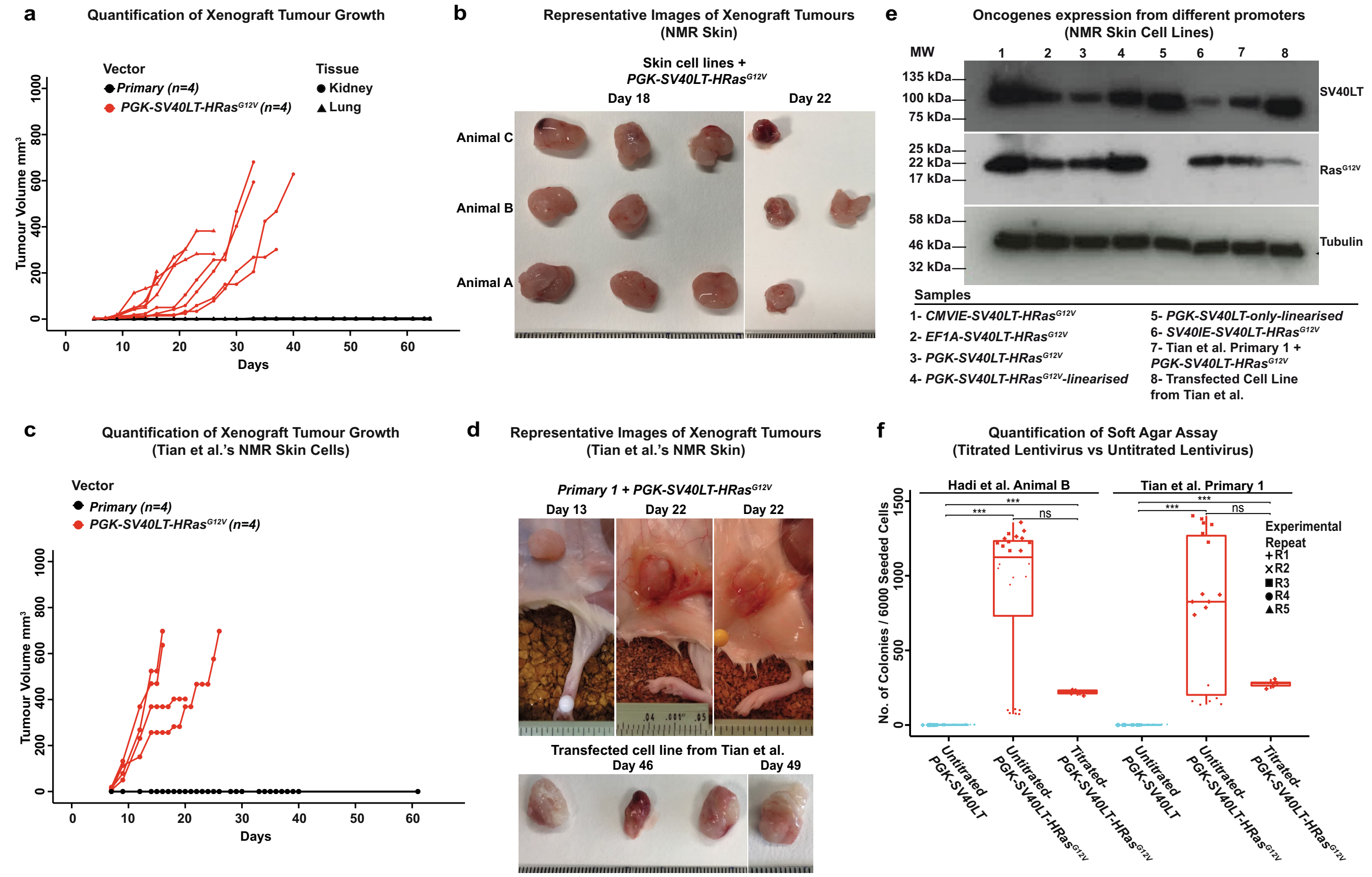
Author contributions F.H. designed and carried out most of the experiments and analysed data. Y.K. helped with cell line generation and performed western blots. K.A.L. and R.U. helped with xenograft experiments. K.B. analysed the RNA-sequencing data. P.B., E.S.J.S. and W.T.K. conceptualized the original ideas for the project. E.S.J.S. and W.T.K. conceived and supervised the study. F.H., E.S.J.S. and W.T.K. wrote manuscript with input from all other authors.

Competing interests The authors declare no competing interests.

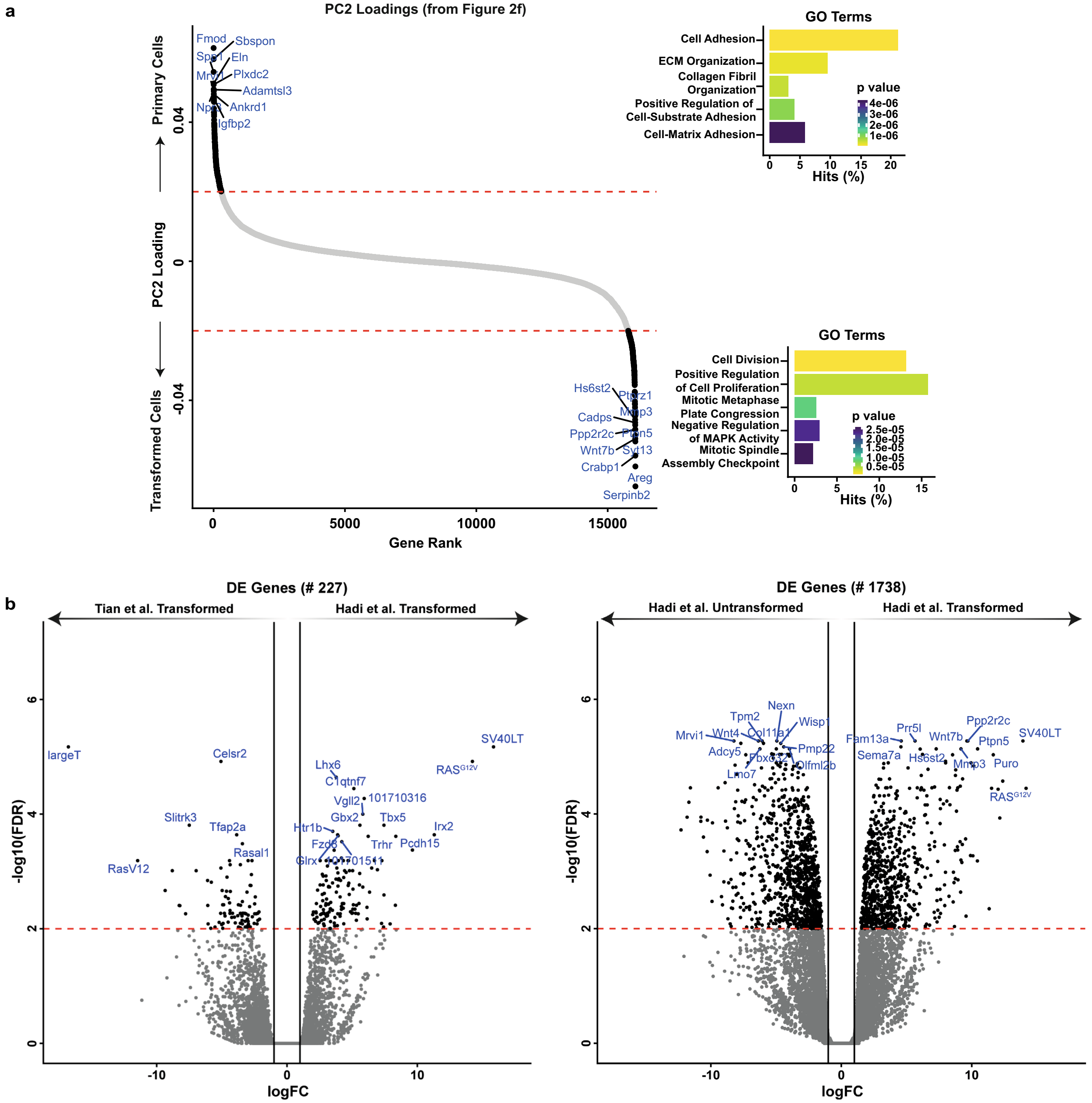
Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2410-x.
Correspondence and requests for materials should be addressed to E.S.J.S. or W.T.K.
Reprints and permissions information is available at http://www.nature.com/reprints.



Extended Data Fig. 1 | NMR cells from different tissues can be transformed. **a**, Schematic of the lentiviral vectors generated and used in this study. CMV, cytomegalovirus promoter; RU5, 5′ long-terminal repeat lacking the U3 region; hU6, human U6 promoter; iScaffold, improved guide RNA (gRNA) scaffold; PGK, mouse *Pgk1* promoter; EF1α, human *EF1A* promoter; SV40E, SV40 early region promoter; CMVIE, CMV immediate early promoter; Puro, puromycin resistance gene; 2A, *Thosea asigna* self-cleaving 2A peptide; ΔU3RU5, enhancer-deleted 3′ long-terminal repeat. **b**, Schematic representation of the different cell lines generated as part of this study using the vectors shown in **a**. **c**–**f**, Quantification of soft-agar assay colonies from different cell lines generated from NMR cells from the pancreas (**c**), lung (**d**), kidney (**e**) and intestine (**f**). For each organ, cells were derived from four different NMRs, except in the case of intestine, for which cells were derived from two different NMRs. Different shapes represent different experimental repeats. Each experiment was performed up to four times and each experimental repeat included six technical replicates. Each data point represents the number of colonies observed from an individual technical replicate. In total, more than 3,300 fields of view (non-overlapping images) were analysed for the pancreas, whereas more than 3,200, 3,800 and 1,300 fields of view were analysed for the lung, kidney and intestine, respectively. It is worth noting that kidney cells can be transformed with *SV40LT* alone as shown in **e**. **c**–**f**, Data were analysed using Wilcoxon rank-sum tests; ***$P \leq 0.0001$; NS, not significant. Box plots are shown as follows: centre line, median; box limits, upper and lower quartiles; whiskers, 1.5× the interquartile range.



Extended Data Fig. 2 | Transformed NMR cells form tumours in NSG mice. a, Quantification of xenograft tumour growth in NSG mice injected with primary or *Pgk1-SV40LT;HRAS*^{G12V}-transduced kidney and lung cells. Each cell line was injected into four mice; each mouse is represented by a single line. Colours represent different vectors and shapes represent different tissues. **b**, Representative images of xenograft tumours shown in Fig. 1b. **c**, Quantification of xenograft tumour growth in NSG mice injected with primary or *Pgk1-SV40LT;HRAS*^{G12V}-transduced NMR skin cells obtained from Tian et al.[5]. Each cell line was injected into four mice; each mouse is represented by a single line. **d**, Representative images of xenograft tumours reported in Fig. 1d. **e**, Western blots showing the expression of SV40LT and HRAS^{G12V} from different promoters in NMR skin cell lines. **f**, Quantification of soft-agar colonies from NMR skin cell lines generated by introduction of *Pgk1-SV40LT* or *Pgk1-SV40LT;HRAS*^{G12V} through lentiviral particles. In each set of experiments, lentiviral particles used for introducing *Pgk1-SV40LT;HRAS*^{G12V} were titrated to keep the number of lentiviral particles transducing each cell to around 1. Each experiment was repeated up to five times (represented by different shapes) and each experimental repeat included six technical replicates. Each data point represents the number of colonies observed from a single technical replicate. Data were analysed using Wilcoxon rank-sum tests; ***$P \leq 0.0001$; NS, not significant. Box plots are shown as follows: centre line, median; box limits, upper and lower quartiles; whiskers, 1.5× the interquartile range.



Extended Data Fig. 3 | Gene expression analysis of transformed NMR cells. **a**, Loadings of PC2 for all genes used to compute the PCA in Fig. 2f ordered by the loading. The top ten genes with the highest or lowest loadings are highlighted in blue. Gene-set enrichment was performed on all genes with a loading higher or lower than 0.02 or −0.02, respectively. The top 5 Gene Ontology (GO) terms (biological processes) are visualized in the plot. Positive PC2 values were associated with untransformed cells whereas negative PC2 values were associated with transformed cells. **b**, Results of the differential expression analysis of various NMR cell lines. Left, transformed cells generated in this study (Hadi et al.) were compared with transformed cells from Tian et al.[5]. Right, transformed cells generated in this study were compared with untransformed from this study. The top 20 differentially expressed (DE) genes as well as the transgenes present in the samples are highlighted in blue. The dashed line represents an FDR-adjusted threshold of $P = 0.01$. The following transgenes are highlighted in the volcano plots: *Puro*, *SV40LT* and *RAS*^{G12V}, which encode PuroR, SV40LT and HRAS^{G12V}, respectively, in the *Pgk1-SV40LT;HRAS*^{G12V} vector generated in the current study (Extended Data Fig. 1a); largeT and RasV12 encode SV40 large T antigen from pSG5-largeT (Addgene, 9053) and HRAS^{G12V} from pCMV-RasV12 (Clontech, 631924), respectively, and are from Tian et al.[5].

nature research

Corresponding author(s): Walid T. Khaled & Ewan. St. John Smith

Last updated by author(s): Jan 13, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist .

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	FASTQ files were aligned to the reference genome of either Mus Musculus (GRCm38 with GRC38.91 annotation from ENSEMBL) or to the HetGla2.0 reference genome with the GCF_000247695.1 annotation from NCBI. The mapping was performed using the STAR aligner (2.7.1a) with default settings and –quantMode GeneCounts to derive gene level counts. The counts were then normalized with the trimmed mean of M-values (TMM) method to account for differences in sequencing depth and library composition19. The principle component analysis (Supplementary Figure 8) for mouse and naked mole rat cell lines was performed on log-transformed CPM (counts-per-million) values for a set of 16054 published orthologous genes16. Genes with high or low loading for PC2 (>0.02 or < -0.02) were characterized by gene set enrichment analysis based on gene-ontology (GO) terms using topGO with default settings (Alexa, A. & Rahnenfuhrer, J. TopGO: Enrichment Analysis for Gene Ontology. R package version 2.28.0 (2016)). All differential expression tests were performed using edgeR20 by fitting a negative binomial generalised log-linear model with the cell line as covariate. Genes that have a higher log fold change than 1 at an FDR of 0.01 were identified using the 'glmTreat' function.
Data analysis	FASTQ files were aligned to the reference genome of either Mus Musculus (GRCm38 with GRC38.91 annotation from ENSEMBL) or to the HetGla2.0 reference genome with the GCF_000247695.1 annotation from NCBI. The mapping was performed using the STAR aligner (2.7.1a) with default settings and –quantMode GeneCounts to derive gene level counts. The counts were then normalized with the trimmed mean of M-values (TMM) method to account for differences in sequencing depth and library composition19. The principle component analysis (Supplementary Figure 8) for mouse and naked mole rat cell lines was performed on log-transformed CPM (counts-per-million) values for a set of 16054 published orthologous genes16. Genes with high or low loading for PC2 (>0.02 or < -0.02) were characterized by gene set enrichment analysis based on gene-ontology (GO) terms using topGO with default settings (Alexa, A. & Rahnenfuhrer, J. TopGO: Enrichment Analysis for Gene Ontology. R package version 2.28.0 (2016)). All differential expression tests were performed using edgeR20 by fitting a negative binomial generalised log-linear model with the cell line as covariate. Genes that have a higher log fold change than 1 at an FDR of 0.01 were identified using the 'glmTreat' function.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The RNAseq data will be deposited into publicly available data repositories and accession numbers provided in the manuscript. All other other supporting raw data will be available from the authors upon request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The number of naked mole-rats analysed in this study was dependent on age and sex. For the majority of the naked mole-rats we attempted to generate as many cell lines from as many tissues as possible using different vectors and transfection. This led to the generation of 106 cell lines. Full details of all the lines generated are available in Supplementary Table 1.
Data exclusions	NA
Replication	All experiments in this study were repeated multiple times (number of experimental repeats is noted for each figure and in figure legends)
Randomization	Cell lines used in the xenograft assays were assigned random names by one experimenter and handed over to a second experimenter who then injected them into mice. Each cell line was injected into four mice and equal number of mice of both genders were used for each cell line. In some cases, animals from different groups were housed in the same cage. The animal technician who measured the tumour masses were unaware of the grouping and cell line identity.
Blinding	Cell lines used in the xenograft assays were assigned random names by one experimenter and handed over to a second experimenter who then injected them into mice. Each cell line was injected into four mice and equal number of mice of both genders were used for each cell line. In some cases, animals from different groups were housed in the same cage. The animal technician who measured the tumour masses were unaware of the grouping and cell line identity.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study	
☐	☒	Antibodies
☐	☒	Eukaryotic cell lines
☒	☐	Palaeontology
☐	☒	Animals and other organisms
☒	☐	Human research participants
☒	☐	Clinical data

Methods

n/a	Involved in the study	
☒	☐	ChIP-seq
☒	☐	Flow cytometry
☒	☐	MRI-based neuroimaging

Antibodies

Antibodies used	The following antibodies were used: SV40LT (sc-148, Santa Cruz #sc-58665), RasG12V (D2H12, Cell Signalling #14412), Tubulin (DM1A, Abcam #ab7291). For details of dilutions of antibodies used please refer to the 'Protein Extraction, Antibodies and Immunoblotting' section of the Material & Methods

Validation	Protein extract from primary cell lines (not expressing any of the transgenes) were used to validate the antibodies.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	106 cell lines were generated from 5 tissues (skin, pancreas, lungs, kidneys and intestines) of 11 different Naked Mole-Rats (Supplementary Table 1). Details of how the cell lines were generated are in the methods section of the manuscript
Authentication	NA
Mycoplasma contamination	All cell lines used in this study were tested and negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	NA

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Cells were derived from 11 (2 females, 9 males) adult naked mole-rats (age range: 37 - 244 weeks). Mice used for xenograft assay were of the NSG strain. Mouse skin cell lines were generated from 23 weeks old C57BL/6 female mice.
Wild animals	This study did not involve the use of wild animals
Field-collected samples	NMRs used in the study were housed in custom-made caging system with conventional mouse/rat cages connected by different lengths of tunnel. Bedding and nesting material were provided along with a running wheel. The room was warmed to 28 °C, with a heat cable to provide extra warmth running under 2-3 cages, and red lighting (08:00 - 16:00) was used.
Ethics oversight	All animal experiments were carried out under Home Office licenses P7EBFC1B1 and P1958D980.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Reply to: Transformation of naked mole-rat cells

https://doi.org/10.1038/s41586-020-2411-9

Jing Zhao[1,2,7], Xiao Tian[3,7], Yabing Zhu[1,7], Zhihui Zhang[3], Elena Rydkina[3], Yongxian Yuan[1], Hongyun Zhang[1], Bhaskar Roy[1], Adam Cornwell[3], Eviatar Nevo[4], Xiaoxiao Shang[5], Runyue Huang[5], Karsten Kristiansen[1,2], Andrei Seluanov[3]✉, Xiaodong Fang[1,5,6]✉ & Vera Gorbunova[3]✉

REPLYING TO **F. Hadi et al. *Nature* https://doi.org/10.1038/s41586-020-2410-x (2020)**

It has been independently demonstrated by us[1] and others[2] that naked mole-rat (NMR) cells are more resistant to *SV40LT* and *HRAS*^{G12V}-induced transformation than mouse cells. In the accompanying Comment, Hadi et al.[3] argue that NMR cells and mouse cells are equally susceptible to oncogenic transformation by *SV40LT* and *HRAS*^{G12V}; however, their observations are based on much higher expression levels of HRAS(G12V) than in the previous studies[1,2]. Here we show, using new RNA-sequencing (RNA-seq) data, that NMR cells are considerably more resistant to transcriptomic changes induced by oncogenic HRAS than mouse, blind mole-rat and human cells, indicating that suppressed RAS signalling is an anti-cancer mechanism in NMR cells that can be interrupted by high expression of HRAS(G12V), rendering NMR cells susceptible to oncogenic transformation. Our results explain that the ostensibly equal susceptibility of NMR and mouse cells to transformation observed by Hadi et al.[3] resulted from high expression levels of HRAS(G12V) overriding the anti-cancer mechanisms of the NMRs.

The key difference between the tumorigenesis experiments by us[1] and Hadi et al.[3] was the strength of the promoters used to drive oncogenic HRAS expression. Here, we clarify that, in our original publication[1], the stable mouse and NMR cell lines used for tumour xenograft experiments (figure 4b of the previous study[1]) were generated by sequentially integrating a NotI-linearized pBabe-puro-largeTcDNA plasmid (Addgene, 14088) followed by puromycin selection and a NotI-linearized pWZL-hygro-HRAS(G12V) plasmid (Addgene, 18749) followed by hygromycin selection. Both *SV40LT* and *HRAS*^{G12V} in these plasmids are driven by a retroviral long terminal repeat (*LTR*) promoter, the same promoter that was used by Liang et al.[2] and numerous other tumorigenesis studies[4–7]. The pSG5-largeT and pCMV-RasV12 plasmids cited by Hadi et al.[3] were used only for anchorage-independent soft-agar growth assays[1], and no stable cell lines were generated with using these plasmids.

By contrast, Hadi et al.[3] used much stronger *CMVIE*, *SV40E*, *Pgk1* and *EF1A* promoters, which drove much higher expression of HRAS(G12V) (extended data figure 2e of Hadi et al.[3]). This prompted us to investigate whether differential levels of oncogene expression explain the discrepancy of observations between our study[1] and that of Hadi et al.[3].

We performed RNA-seq analysis on the stable mouse and NMR cell lines generated in our original study[1]. Stable human and blind mole-rat cell lines were also included to test whether NMR cells have a unique response to oncogenes compared with other long-lived species. All transgenes used to generate stable cell lines were driven by *LTR* promoters (see Supplementary Methods). The tumorigenicity of the cell lines from each species[1,8] is summarized in Extended Data Fig. 1a. Similar numbers of genes were detected in all samples by RNA-seq analysis (Extended Data Table 1). Orthologue identification revealed that 13,276 orthologues were shared by all four species, the expression of which was used for further analysis.

Principal component analysis (PCA) showed that the segregation patterns induced by oncogenes were different across the four species (Extended Data Fig. 1b–e). Notably, in NMR cells, separation of samples in PCA plots was primarily driven by SV40LT, but not by HRAS(G12V) (Extended Data Fig. 1d), indicating that NMR cells are refractory to the transcriptomic changes induced by HRAS(G12V). Furthermore, the number of SV40LT- and HRAS(G12V)-induced differentially expressed genes (DEGs) was similar in mouse, blind mole-rat and human cells, but was much lower in NMR cells (Fig. 1a). Consistently, multiple Gene Ontology (GO) terms that were significantly changed by SV40LT and HRAS(G12V) in mouse cells, such as cell cycle, cell division and mitotic nuclear division, were not altered or altered to a much lower extent in NMR cells (Extended Data Fig. 2a).

We next analysed the effect of SV40LT and HRAS(G12V) expression on transcriptomic changes separately. Notably, SV40LT expression induced more DEGs in NMR cells than in the other three species (Fig. 1b and Extended Data Fig. 2b). By contrast, HRAS(G12V) expression induced much fewer DEGs (Fig. 1c) and fewer changes in GO terms (Extended Data Fig. 2c) in the NMR cells, supporting a refractory response of NMR cells to *HRAS*^{G12V} expression as seen in the PCA. The levels of HRAS(G12V) expression were similar across species (Fig. 1d), excluding the possibility that the refractory response of NMR cells to HRAS(G12V) was due to differential HRAS(G12V) expression.

It was previously demonstrated that the expression level of HRAS(G12V) determines human cell transformation[9]. We next asked whether a high level of HRAS(G12V) expression in NMR cells could lead to transformation and explain the discrepancy in the observations between our study and that of Hadi et al.[3]. *HRAS*^{G12V} driven by a *CAG* promoter was stably integrated into *SV40LT*-expressing NMR cells, generating *LTR-SV40LT;CAG-HRAS*^{G12V} NMR cells (see Methods). The generated cell lines expressed a much higher level of HRAS(G12V) than the *LTR*-driven *SV40LT;HRAS*^{G12V} (hereafter, *LTR-SV40LT;HRAS*^{G12V}) NMR cells (Fig. 2a). We next tested the tumorigenicity of the stable cell lines. The tumour incidence of *LTR-SV40LT;HRAS*^{G12V} NMR cells was relatively low (around 29%, 4 tumours formed out of 14 xenografts; Fig. 2b) compared with mouse *LTR-SV40LT;HRAS*^{G12V} cells (100%, 6 tumours formed out of 6 xenografts), confirming our original observation that

---

[1]BGI Genomics, BGI-Shenzhen, Shenzhen, China. [2]Laboratory of Genomics and Molecular Biomedicine, Department of Biology, University of Copenhagen, Copenhagen, Denmark. [3]Department of Biology, University of Rochester, Rochester, NY, USA. [4]Institute of Evolution, University of Haifa, Haifa, Israel. [5]The Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, China. [6]The Third Xiangya Hospital of Central South University, Changsha, China. [7]These authors contributed equally: Jing Zhao, Xiao Tian, Yabing Zhu. ✉e-mail: andrei.seluanov@rochester.edu; fangxd@bgi.com; vera.gorbunova@rochester.edu

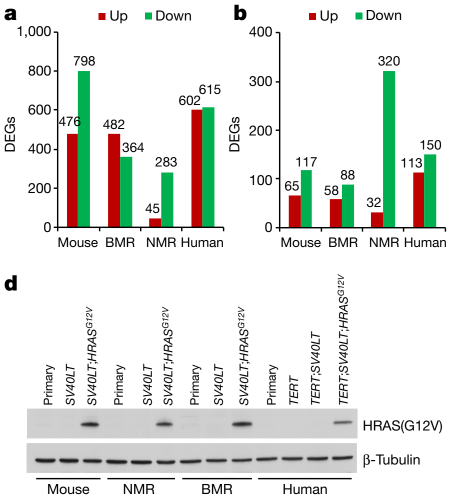


Fig. 1 | Differential response of NMR cells to HRAS(G12V) expression. **a**, The number of DEGs induced by the combination of *SV40LT* and *HRAS*^{G12V} in *SV40LT;HRAS*^{G12V} cells compared with primary cells across species. BMR, blind mole rat. **b**, The number of DEGs induced by *SV40LT* in *SV40LT* cells compared with primary cells. **c**, The number of DEGs induced by *HRAS*^{G12V} in cells expressing both *HRAS*^{G12V} and *SV40LT*, compared with cells expressing *SV40LT*. **d**, Western blot of the stable cell lines used for RNA-seq. All transgenes are driven by *LTR* promoters. $n=3$ per cell line for RNA-seq analyses. Western blots were performed at least three times and a representative image is shown.

NMR cells are more resistant to oncogenic transformation than mouse cells[1]. However, NMR cells with a higher expression of HRAS(G12V) driven by the *CAG* promoter (*LTR-SV40LT;CAG-HRAS*^{G12V} NMR cells) formed tumours 100% of the time (18 tumours formed out of 18 xenografts) and displayed a significantly reduced tumour latency (Fig. 2b, c). Tumorigenicity of *LTR-SV40LT;HRAS*^{G12V} NMR cells could also be increased by high expression of SV40 small T (*SV40ST*) driven by a *CAG* promoter (*LTR-SV40LT;HRAS*^{G12V}*;CAG-SV40ST* NMR cells) (around 75%, 9 tumours formed out of 12 xenografts), but not by lower expression of *SV40ST* driven by a *LTR* promoter (*LTR-SV40LT;HRAS*^{G12V}*;SV40ST* NMR cells) (Fig. 2b, c). NMR cells expressing SV40LT and SV40ST (*CAG-SV40LT;SV40ST* NMR cells) did not form tumours (Fig. 2b, c). The combination of *SV40LT*, *SV40ST*, *TERT* and *HRAS*^{G12V} driven by *CAG* promoters efficiently transformed human skin fibroblasts (Fig. 2b, c), confirming previous publications[4,5,10].

To decipher the molecular mechanisms that underlie the differential response to high and moderate levels of HRAS(G12V) expression in NMR cells, we thoroughly analysed the Ras effector pathways, including ERK and AKT, and identified inherently decreased PI3K–AKT signalling downstream of RAS[11]. Furthermore, AKT activity was derepressed by high expression of HRAS(G12V)[11]. These results indicate that decreased PI3K–AKT signalling is an evolutionary adaption to resist oncogenic transformation in NMR cells. The PI3K–AKT signalling pathway has also been recognized to regulate longevity. In *Caenorhabditis elegans* and *Drosophila*, abolishing PI3K–AKT signalling significantly extends lifespan[12,13]. In mice, partial inactivation of PI3K or AKT enhances metabolic function and extends lifespan[14,15]. Therefore, the natural suppression of the PI3K–AKT pathway in NMR probably contributes not only to cancer resistance, but also to a long lifespan.

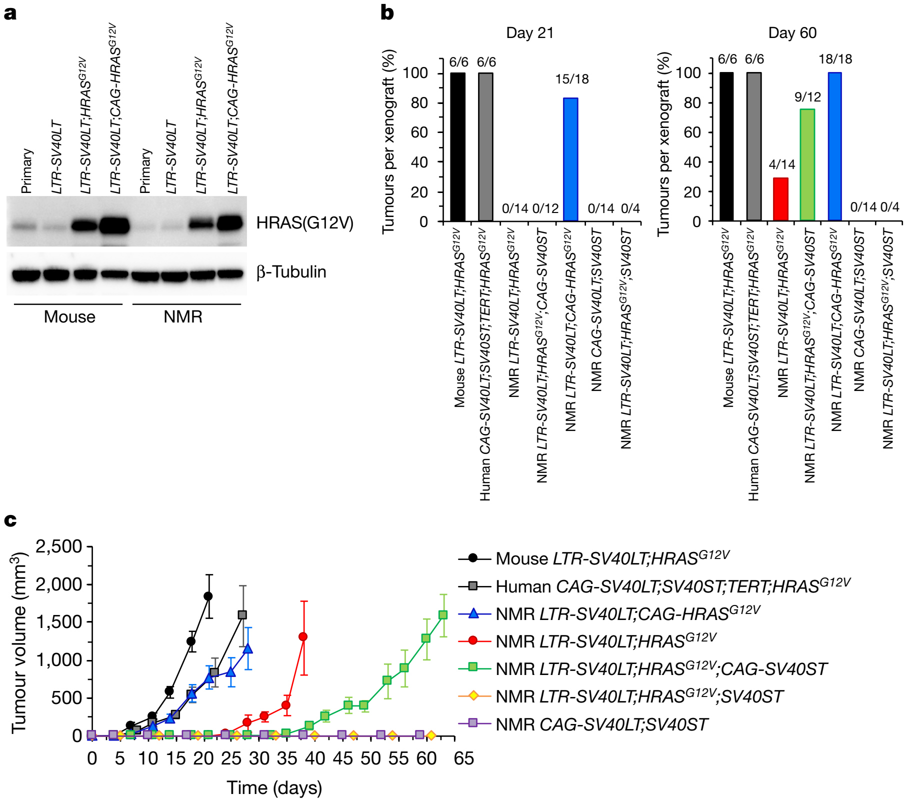


Fig. 2 | HRAS(G12V) expression levels determine oncogenic transformation in NMR cells. **a**, Western blot analysis of mouse and NMR cells that express low and high levels of HRAS(G12V). Western blots were performed at least three times and a representative image is shown. **b**, Summary of tumour formation in xenografts of mouse, human and NMR cells expressing oncoproteins under the control of different promotors. *LTR* is a moderate promoter that has commonly been used to transform mouse and human cells (Supplementary Methods). *CAG* is a strong promoter. The fractions indicate the number of tumours formed out of the number of injections. **c**, Growth curves of tumours formed by the transformed cells. $n=$ the number of tumours as indicated in **b**. Data are mean ± s.e.m. The order of the transgenes indicates the order of integration and selection in the cells.

Hadi et al.[3] also observed a lower clonogenic efficiency and longer tumour latency using the stable *LTR-SV40LT;HRAS*^{G12V} NMR cells that we provided. Both the low frequency and long latency of tumour formation of the *LTR-SV40LT;HRAS*^{G12V} NMR cells suggest that additional genetic changes may have accumulated in a small fraction of the cell population during passaging after the introduction of *SV40LT* and *HRAS*^{G12V}, which overcame the tumour-suppressing mechanisms of NMR cells, such as the suppressed RAS signalling pathway and the hyaluronan barrier. The results by Hadi et al.[3] using our transformed cells support our conclusion that NMR cells are more resistant to transformation than mouse cells, but this difference could only be revealed with a moderate level of oncogene expression. Very high levels of oncogenic RAS expression are probably not physiologically relevant as indicated by a previous study[9]. The naturally decreased RAS signalling pathway in NMRs makes the expression levels of RAS particularly important for cross-species comparisons; artificially high RAS expression levels override the natural cancer-resistance mechanisms in NMRs.

Methods

Methods are available in the Supplementary Information.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The RNA-seq data have been deposited in the NCBI Sequence Read Archive (SRA) (SRP133455). Uncropped western blots are provided as Supplementary Fig. 1. All other data that support the findings of this study can be found as in bioRxiv preprint[11] or are available from the corresponding authors upon reasonable request.

Code availability

Source code are available upon request.

1. Tian, X. et al. High-molecular-mass hyaluronan mediates the cancer resistance of the naked mole rat. *Nature* **499**, 346–349 (2013).
2. Liang, S., Mele, J., Wu, Y., Buffenstein, R. & Hornsby, P. J. Resistance to experimental tumorigenesis in cells of a long-lived mammal, the naked mole-rat (*Heterocephalus glaber*). *Aging Cell* **9**, 626–635 (2010).
3. Hadi, F. et al. Transformation of naked mole-rat cells. *Nature* https://doi.org/10.1038/s41586-020-2410-x (2020).
4. Hahn, W. C. et al. Enumeration of the simian virus 40 early region elements necessary for human cell transformation. *Mol. Cell. Biol.* **22**, 2111–2123 (2002).
5. Rangarajan, A., Hong, S. J., Gifford, A. & Weinberg, R. A. Species- and cell type-specific requirements for cellular transformation. *Cancer Cell* **6**, 171–183 (2004).
6. Zhao, J. J. et al. Human mammary epithelial cell transformation through the activation of phosphatidylinositol 3-kinase. *Cancer Cell* **3**, 483–495 (2003).
7. Serrano, M., Lin, A. W., McCurrach, M. E., Beach, D. & Lowe, S. W. Oncogenic *ras* provokes premature cell senescence associated with accumulation of p53 and p16INK4a. *Cell* **88**, 593–602 (1997).
8. Tian, X. et al. Evolution of telomere maintenance and tumour suppressor mechanisms across mammals. *Phil. Trans. R. Soc. Lond. B* **373**, 20160443 (2018).
9. Elenbaas, B. et al. Human breast cancer cells generated by oncogenic transformation of primary mammary epithelial cells. *Genes Dev*. **15**, 50–65 (2001).
10. Hahn, W. C. et al. Creation of human tumour cells with defined genetic elements. *Nature* **400**, 464–468 (1999).
11. Zhao, J. *et al*. Dampened PI3K/AKT signaling contributes to cancer resistance of the naked mole rat. Preprint at bioRxiv https://doi.org/10.1101/2020.02.27.967729 (2020).
12. Ayyadevara, S., Alla, R., Thaden, J. J. & Shmookler Reis, R. J. Remarkable longevity and stress resistance of nematode PI3K-null mutants. *Aging Cell* **7**, 13–22 (2008).
13. Moskalev, A. A. & Shaposhnikov, M. V. Pharmacological inhibition of phosphoinositide 3 and TOR kinases improves survival of *Drosophila melanogaster*. *Rejuvenation Res*. **13**, 246–247 (2010).
14. Nojima, A. et al. Haploinsufficiency of Akt1 prolongs the lifespan of mice. *PLoS One* **8**, e69178 (2013).
15. Foukas, L. C. et al. Long-term p110α PI3K inactivation exerts a beneficial effect on metabolism. *EMBO Mol. Med*. **5**, 563–571 (2013).

Acknowledgements This study is supported by US National Institutes of Health grant AG047200 to V.G. and A.S., and Natural Science Foundation of China grant 81672818, Science Technology and Innovation Committee of Shenzhen Municipality grant JCYJ20160331190123578 and Guangzhou Science and Technology Program key project 201604020005 to X.F.

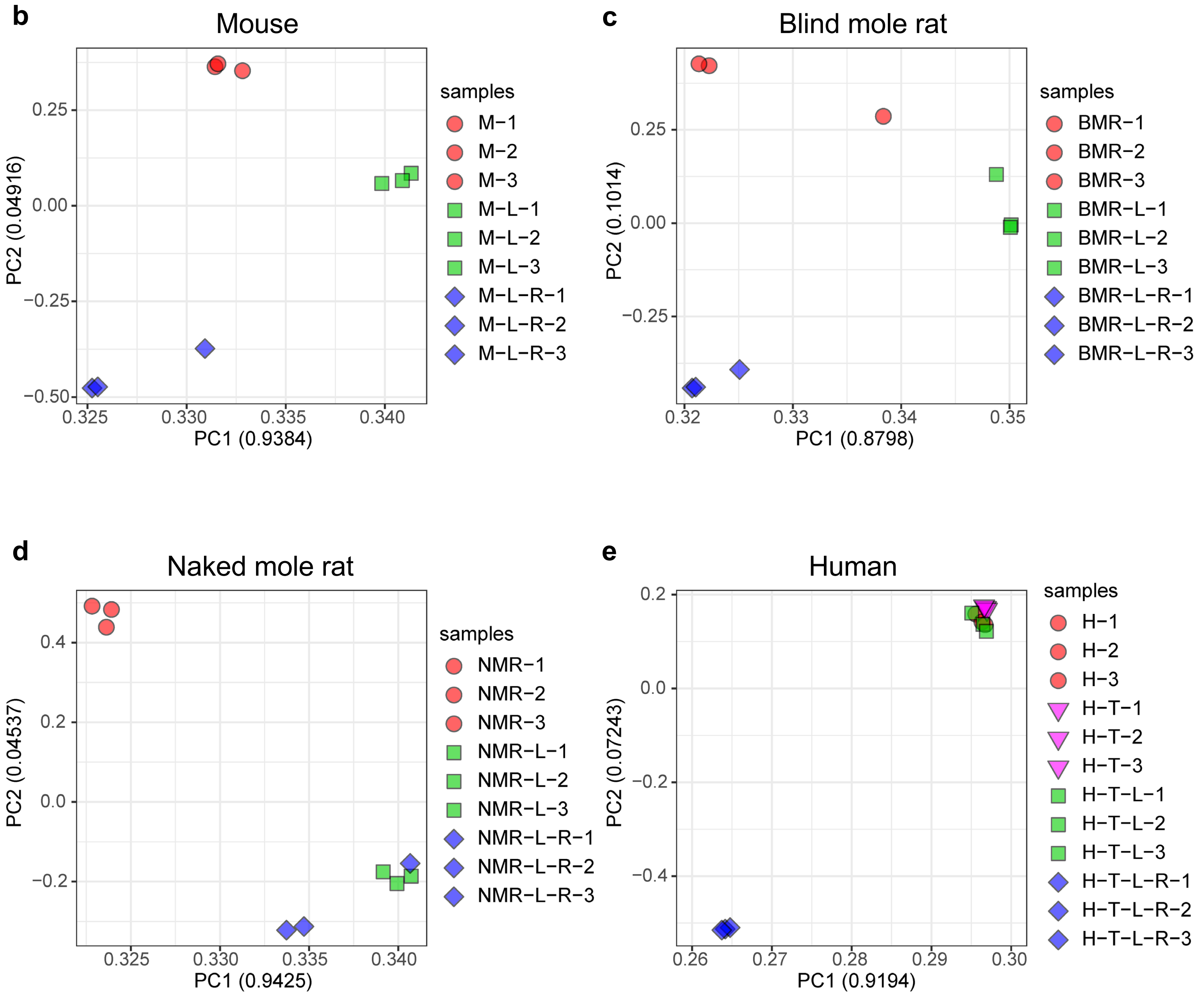
Author contributions The list of authors is different from the original study, because new data were included in this Reply. J.Z., X.T., Y.Z., X.F. and V.G. designed research. X.T., Z.Z. and E.R. generated the stable cell lines and performed western blot and xenograft experiments. H.Z. constructed libraries and sequenced the samples. J.Z., Y.Z., Y.Y., B.R., A.C., X.S. and R.H. performed data control and bioinformatics analysis. E.N. provided samples. K.K. contributed to the result interpretation and paper writing. A.S., X.F., and V.G. supervised the study. J.Z., X.T., Y.Z., A.S., X.F. and V.G. wrote the paper with input from all authors.

Competing interests The authors declare no competing interests.

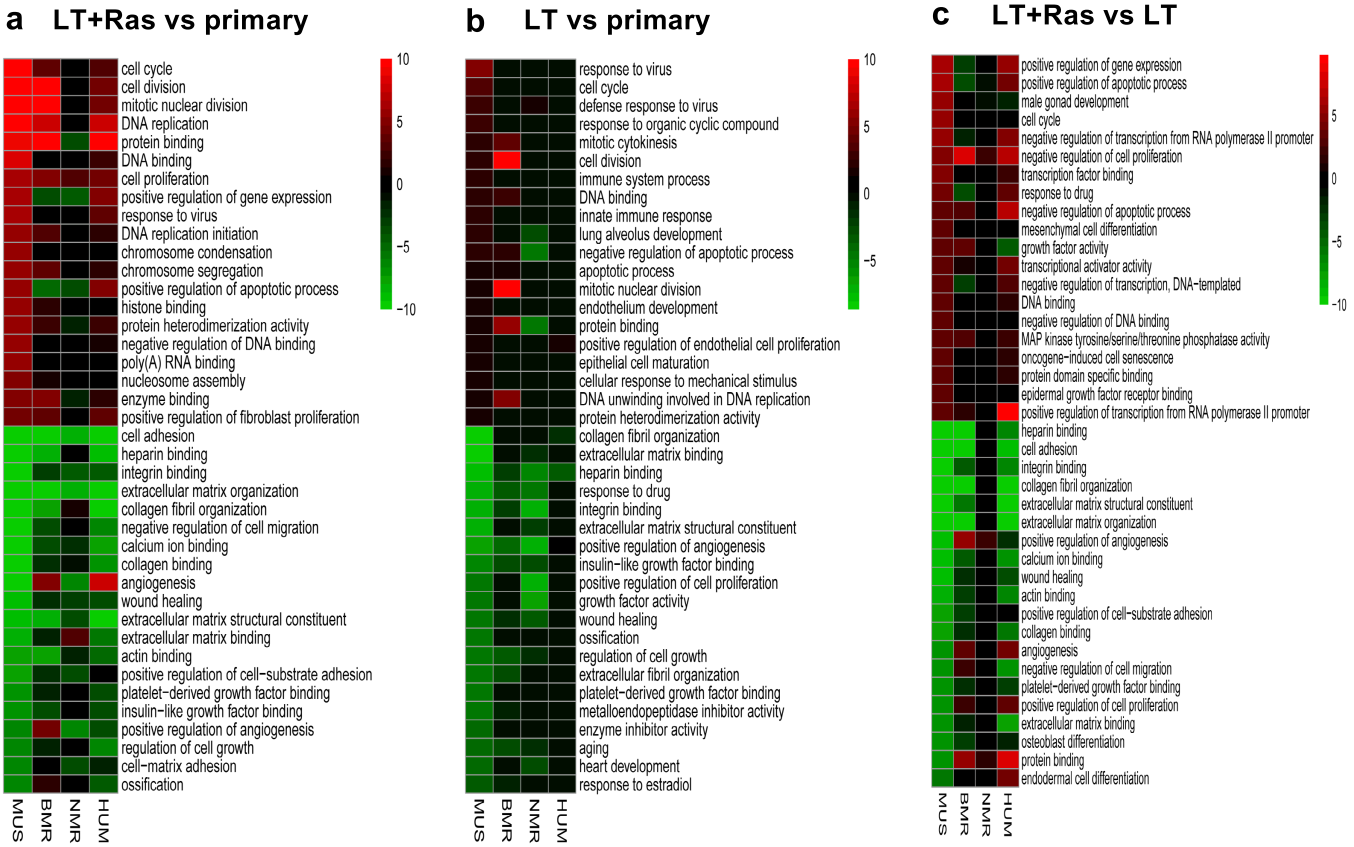
Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2411-9.
Correspondence and requests for materials should be addressed to A.S., X.F. or V.G.
Reprints and permissions information is available at http://www.nature.com/reprints.
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

a

	Primary	LT	LT + HRasV12	*TERT*	*TERT* + LT	*TERT* + LT + HRasV12
Mouse	No	No	Yes	–	–	–
BMR	No	No	Yes	–	–	–
NMR	No	No	Low	–	–	–
Human	No	–	–	No	No	No



Extended Data Fig. 1 | Tumorigenicity of the cell lines and PCA of the RNA-seq datasets. a, Tumorigenicity of the cell lines. Cells not analysed in this study are indicated with a '-'. All transforming genes were driven by a retroviral *LTR* promoter. **b**–**e**, PCA of the RNA-seq datasets. PCA was done using the expression of 13,276 orthologous genes across all four species (mouse (**b**), blind mole rat (**c**), NMR (**d**) and human (**e**)). The first two principal components of each analysis were extracted. Values in parentheses indicate the variance explained by each of the principal components. L, SV40LT; R, HRAS(G12V); T, human TERT. All transgenes are driven by *LTR* promoters.



Extended Data Fig. 2 | Enrichment analysis of DEGs in the transformed cells across species. a, GO enrichment analysis of the DEGs derived from Fig. 1a. **b**, GO enrichment analysis of the DEGs derived from Fig. 1b. **c**, GO enrichment analysis of the DEGs derived from Fig. 1c. Biological process and molecular function terms are shown. All GO terms in **a**–**c** are arranged in the order of decreasing significance in the mouse cells. The top 20 significantly upregulated (red) and downregulated (green) terms are shown. LT, SV40LT; RAS, HRAS(G12V). All transgenes are driven by *LTR* promoters.

Extended Data Table 1 | Statistics for RNA-seq data and expressed genes for all samples

Sample	SRA accession number	Number of Reads	Number of bases	Number of genes
BMR-1	SRX3741501	34943168	5241475200	12816
BMR-2	SRX3741500	34237168	5135575200	12913
BMR-3	SRX3741499	34763768	5214565200	12783
BMR-L-1	SRX3741502	33980456	5097068400	12983
BMR-L-2	SRX3741503	34715316	5207297400	13185
BMR-L-3	SRX3741504	33697198	5054579700	13113
BMR-L-R-1	SRX3741505	34896340	5234451000	13370
BMR-L-R-2	SRX3741506	33554738	5033210700	13352
BMR-L-R-3	SRX3741507	35027026	5254053900	13428
HCA2-1	SRX3741508, SRX3741480	34428058	5164208700	13549
HCA2-2	SRX3741481, SRX3741478	34559592	5183938800	13533
HCA2-3	SRX3741484, SRX3741479	34311928	5146789200	13570
HCA2-T-1	SRX3741485, SRX3741482	34369286	5155392900	13537
HCA2-T-2	SRX3741483, SRX3741472	34482134	5172320100	13531
HCA2-T-3	SRX3741517, SRX3741473	34939600	5240940000	13546
HCA2-T-L-1	SRX3741516, SRX3741515	34654506	5198175900	13580
HCA2-T-L-2	SRX3741514, SRX3741513	34653804	5198070600	13517
HCA2-T-L-3	SRX3741512, SRX3741511	35401832	5310274800	13594
HCA2-T-L-R-1	SRX3741519, SRX3741510	34482018	5172302700	13755
HCA2-T-L-R-2	SRX3741518, SRX3741470	34917538	5237630700	13780
HCA2-T-L-R-3	SRX3741471, SRX3741469	35351198	5302679700	13771
M-1	SRX3741509	34856638	5228495700	13364
M-2	SRX3741474	33703474	5055521100	13165
M-3	SRX3741475	34872236	5230835400	13376
M-L-1	SRX3741476	34876306	5231445900	13423
M-L-2	SRX3741477	33989634	5098445100	13457
M-L-3	SRX3741496	34168260	5125239000	13452
M-L-R-1	SRX3741498	34267706	5140155900	13229
M-L-R-2	SRX3741462	33990568	5098585200	13355
M-L-R-3	SRX3741461	34491972	5173795800	13195
NMR-1	SRX3741464	34807890	5221183500	13144
NMR-2	SRX3741463	34095594	5114339100	13148
NMR-3	SRX3741458	34153410	5123011500	13339
NMR-L-1	SRX3741460, SRX3741457	35874752	5381212800	13343
NMR-L-2	SRX3741466, SRX3741459	36027014	5404052100	13436
NMR-L-3	SRX3741492, SRX3741465	35718470	5357770500	13410
NMR-L-R-1	SRX3741493, SRX3741490	35318112	5297716800	13443
NMR-L-R-2	SRX3741491, SRX3741488	35524056	5328608400	13510
NMR-L-R-3	SRX3741489, SRX3741486	34950234	5242535100	13508

nature research

Corresponding author(s):	Andrei Seluanov, Xiaodong Fang, Vera Gorbunova
Last updated by author(s):	Apr 3, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist .

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	RNA-seq: Illumina HiSeq 4000 qRT-PCR: Bio-Rad CFX manager
Data analysis	Data analysis was performed using R version v3.1.1. Key packages used were: SOAPnuke (V1.5), Cufflinks (v2.2.1), samtools (v0.1.19), HISAT2 (v2.0.4), DESeq2 (v1.4.5), topGO (v2.16.0). DAVID (v6.8) was used for enrichment analysis.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The RNA-seq data has been deposited in NCBI's Sequence Read Archive (SRA) and is accessible through the accession number SRP133455. Uncropped Western blots are included in Extended Data Fig. 3. All other relevant data that support the findings of this study are either within the article or available from the corresponding author upon reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical analysis was performed to predetermine sample size. The sample size was chosen based on standard practices in the field. Primary fibroblasts were isolated from at least 6 individual animals of each rodent species.
Data exclusions	No data was excluded.
Replication	All attempts at replication were successful.
Randomization	Primary fibroblasts were isolated from randomized animals of the mouse, blind mole-rat, and naked mole-rat colonies. The NIH-III nude mice used for xenograft experiment were randomized.
Blinding	For exnograft experiment, researchers monitoring the mice and measuring the tumor size were blinded to treatment conditions. No blinding was performed for RNA-seq analysis.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study	
☐	☒	Antibodies
☐	☒	Eukaryotic cell lines
☒	☐	Palaeontology
☐	☒	Animals and other organisms
☒	☐	Human research participants
☒	☐	Clinical data

Methods

n/a	Involved in the study	
☒	☐	ChIP-seq
☒	☐	Flow cytometry
☒	☐	MRI-based neuroimaging

Antibodies

Antibodies used	H-Ras V12 (ab140962, Abcam), β-tubulin (ab6046, Abcam).
Validation	Both antibodies have been widely cited in the literature. Antibody specificity and quality validation were performed by the manufacturers (see manufacturers' webpages for further information). Antibodies were further validated by overexpression controls or expected molecular weight of the targets.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Primary fibroblasts from mice, naked mole rats, and blind mole rats were isolated from underarm skin samples of at least 6 individual animals of each species. The mouse samples were from C57BL/6 mice. The naked mole-rat and the blind mole-rat skin samples are from the University of Rochester colonies. The human skin fibroblasts were a gift from Pereira-Smith lab at University of Texas Health Science Center at San Antonio.
Authentication	Primary fibroblasts were authenticated by their distinctive cell morphology and expected growth rate based on over 10 years' experience of cell cultures of different rodent species. Blind mole-rat and Naked mole-rat cells were further verified by the viscosity of cell culture media due to the presence of high molecular weight hyaluronan.
Mycoplasma contamination	All cell lines were tested routinely and confirmed negative for mycoplasma contamination.

Commonly misidentified lines (See ICLAC register)	Not used

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	C57BL6/J wild type mice were purchased from Jackson Laboratory. The naked mole-rat and the blind mole-rat skin samples are from the University of Rochester colonies. Both sexes and both young and older animals were used for all three rodent species. Two- to three-month-old female NIH-III nude mice (Crl:NIH-Lystbg-JFoxn1nuBtkxid) were purchased from Charles River Laboratories and used to establish xenografts.
Wild animals	Not used.
Field-collected samples	Not used.
Ethics oversight	All animal experiments were approved and performed in accordance with the guidelines set up by the University of Rochester Committee on Animal Resources.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

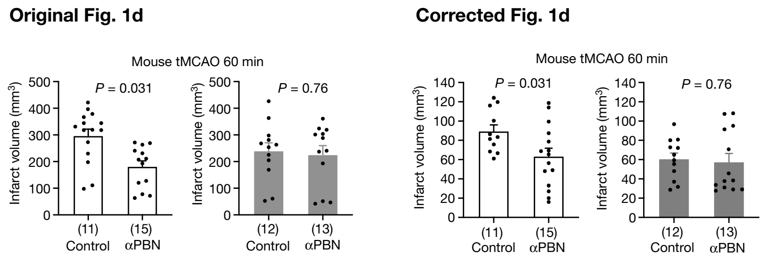
Author Correction: Potential circadian effects on translational failure for neuroprotection

https://doi.org/10.1038/s41586-020-2427-1

Correction to: *Nature* https://doi.org/10.1038/s41586-020-2348-z

Elga Esposito, Wenlu Li, Emiri T. Mandeville, Ji-Hyun Park, Ikbal Şencan, Shuzhen Guo, Jingfei Shi, Jing Lan, Janice Lee, Kazuhide Hayakawa, Sava Sakadžić, Xunming Ji & Eng H. Lo

In this Article, the graphs in Fig. 1d were inadvertently duplicated from Fig. 1b. Figure 1 of this Amendment shows the corrected Fig. 1d alongside the incorrect, published Fig. 1d, for transparency to readers. The original Article has been corrected.



Original Fig. 1d

Corrected Fig. 1d

Fig. 1 | This figure displays the corrected and the incorrect published Fig. 1d of the original Article.

Author Correction: A naturally occurring antiviral ribonucleotide encoded by the human genome

https://doi.org/10.1038/s41586-020-2322-9

Correction to: *Nature* https://doi.org/10.1038/s41586-018-0238-4

Anthony S. Gizzi, Tyler L. Grove, Jamie J. Arnold, Joyce Jose, Rohit K. Jangra, Scott J. Garforth, Quan Du, Sean M. Cahill, Natalya G. Dulyaninova, James D. Love, Kartik Chandran, Anne R. Bresnick, Craig E. Cameron & Steven C. Almo

In this Letter, the isotopologue used to demonstrate hydrogen atom abstraction from CTP was incorrectly identified as 2′,3′,4′,5′,5-deutero-CTP, whereas the isotopologue we used was 3′,4′,5′,5-deutero-CTP. (Boldface indicates where changes were made.) In the main text, the sentence “Incubation of rVIP with SAM and CTP deuterated at the **2**′, 3′, 4′, 5′ and 5 positions (deuCTP), increased the −*m*/*z* of 5′-dA from 250.1 to 251.1, consistent with the transfer of one deuterium from deuCTP to 5′-dA.” should have read “Incubation of rVIP with SAM and CTP deuterated at the 3′, 4′, 5′ and 5 positions (deuCTP) increased the −*m*/*z* of 5′-dA from 250.1 to 251.1, consistent with the transfer of one deuterium from deuCTP to 5′-dA.”. In the Extended Data Fig. 3 legend the sentence “When rVIP was incubated with SAM and CTP deuterated at the **2**′, 3′, 4′, 5′ and 5 positions (deuCTP), the −*m*/*z* of 5′-dA increased from 250.1 to 251.1, consistent with the transfer of one deuterium from deuCTP to 5′-dA•.” should have read: “When rVIP was incubated with SAM and CTP deuterated at the 3′, 4′, 5′ and 5 positions (deuCTP), the −*m*/*z* of 5′-dA increased from 250.1 to 251.1, consistent with the transfer of one deuterium from deuCTP to 5′-dA•.”

Similarly, the name of the isotopologue appeared incorrectly in the Supplementary Information. In the Supplementary Information section ‘Hydrogen atom abstraction specificity’ the sentence “Reactions of 100 µL total volume contained 100 µM rVIP, 50 mM HEPES pH 7.5, 150 mM KCl, 1 mM SAM, and 1 mM deuterium-labeled CTP (site-specifically labeled CTP contains deuterium at the following positio**n; (2'-^{2}H**, 3'-^{2}H, 4'-^{2}H, or 5'-$^{2}H_2$-CTP)).” should have read: "Reactions of 100 µL total volume contained 100 µM rVIP, 50 mM HEPES pH 7.5, 150 mM KCl, 1 mM SAM, and 1 mM deuterium-labelled CTP (site-specifically labelled CTP contains deuterium at the following positions: 3'-^{2}H,4'-^{2}H,5'-$^{2}H_2$,5-^{2}H-CTP or 2'-^{2}H-CTP,3'-^{2}H-CTP,4'-^{2}H-CTP or 5'-$^{2}H_2$-CTP). In the Supplementary Information section ‘2'-deoxy-CTP vs deutero-CTP substrate specificity assays’ the sentence “Reactions of 100 µL total volume contained 100 µM rVIP, 50 mM HEPES pH 7.5, 150 mM KCl, 1 mM SAM, 1 mM deutero-CTP (**2'-^{2}H**, 3'-^{2}H, 4'-^{2}H, 5'-$^{2}H_2$**, and** 5-^{2}H −CTP), and unlabeled 1 mM 2'-deoxy-CTP.” should have read: “Reactions of 100 µL total volume contained 100 µM rVIP, 50 mM HEPES pH 7.5, 150 mM KCl, 1 mM SAM, 1 mM deutero-CTP (3'-^{2}H,4'-^{2}H,5'-$^{2}H_2$,5-^{2}H-CTP), and unlabelled 1 mM 2'-deoxy-CTP.” These errors have been corrected online.

Publisher Correction: Constraint on the matter–antimatter symmetry-violating phase in neutrino oscillations

https://doi.org/10.1038/s41586-020-2415-5

Correction to: *Nature* https://www.nature.com/articles/s41586-020-2177-0

The T2K Collaboration

In the Methods section of this Article, owing to an error during the production process, some reference citations were incorrectly cited. From "For the initial nuclear state… applying alternative multi-pion production tuning[54]", the citation numbers have been corrected. The reference list is correct. The original Article has been corrected online.

Retraction Note: Fatty acids and cancer-amplified ZDHHC19 promote STAT3 activation through *S*-palmitoylation

https://doi.org/10.1038/s41586-020-2414-6

Retraction to: *Nature* https://doi.org/10.1038/s41586-019-1511-x

Jixiao Niu, Yang Sun, Baoen Chen, Baohui Zheng, Gopala K. Jarugumilli, Sarah R. Walker, Aaron N. Hata, Mari Mino-Kenudson, David A. Frank & Xu Wu

We, the authors, are retracting this Letter owing to anomalies in Figs. 2j, 3a and 3d. We and the *Nature* editors reviewed the data and found that these anomalies were present in the digital scans of the original western blot films, which could not be located for further examination. Although repeat experiments at the time of writing the paper have reproduced the data, the anomalies undermine confidence in the published Letter. We therefore believe that the most appropriate course of action is to retract this Letter in its entirety. We thank the readers who brought the issues to our attention, and apologize to the scientific community. We remain confident that the key findings of the Letter are valid and reproducible. All authors agree with the Retraction.

Work

Your story

Send your careers story to: naturecareerseditor @nature.com

NGUYEN HUY KHAM/REUTERS

As the coronavirus has spread around the world, so has misinformation.

FIGHTING CORONAVIRUS MISINFORMATION

Bogus remedies, myths and fake news about COVID-19 can cost lives. Here's how some scientists are fighting back. **By Nic Fleming**

Eating sea lettuce or injecting disinfectant won't prevent you from getting COVID-19. Holding your breath for ten seconds is not a test for SARS-CoV-2. The rapid global spread of COVID-19 has been accompanied by what the World Health Organization has described as a "massive infodemic". Huge demand for information on the disease, its toll on health-care systems and lives, and the many unanswered questions about a virus that was discovered only in December, have created the perfect breeding ground for myths, fake news and conspiracy theories. Some can be dismissed as ludicrous and largely harmless, but others are life-threatening.

Scientists are well placed to help to hold back the tide of COVID-19 misinformation — but should they get involved in time-consuming, and sometimes bruising, efforts to do so? For those who sign up, how can coronavirus untruths best be confronted? Should scientists restrict interventions to their areas of expertise? Is countering falsehoods about the pandemic purely a public service, or might there be career benefits?

"I think scientists need to get out there on the front line, if they are comfortable doing so," says Jevin West, who is a data scientist at the University of Washington in Seattle. "By countering misinformation about COVID-19, they can help policymakers avoid introducing harmful policies, improve public understanding of the pandemic and, most importantly, save lives."

Among the many changes wrought by COVID-19 is a widespread increase in news consumption. A March survey of 13 countries by market-research company GlobalWebIndex found that, as a result of the pandemic, 67% of those surveyed are watching more news coverage, and that half of that subset are spending significantly more time doing so (see go.nature.com/2yznjku). We're "looking for good news or inside information about COVID-19 because it affects our health, and that of our friends and families," says West. "That makes us more vulnerable to being fooled."

West co-created Calling Bullshit, a course on how to spot and counter false appeals to scientific and statistical evidence (see 'Eight ways to spot misinformation'), and in December, co-founded and became director of his university's new Center for an Informed Public, whose core aims include researching rumours and misinformation during crises. It's been a busy few months for West and his colleagues.

The misinformation world

False medical claims are a key focus for those seeking to minimize potential harms. Researchers at the Taiwan FactCheck Center have, for example, spent a large proportion of their time debunking reports about fake remedies and tests since late January. Examples include claims that smelling sesame and other plant oils, breathing in steam or cleaning the nostrils with salty water can kill SARS-CoV-2 before it reaches the lungs.

Some who share myths are simply misguided, but others are driven by profit. In March, the US Food and Drug Administration warned companies and individuals, including Alex Jones, owner of the fake-news website Infowars, and televangelist Jim Bakker, to stop touting the benefits of unproven COVID-19 treatments such as colloidal silver, which they were selling. Another way to profit from fake news is advertising revenue. "About half of the disinformation we see is about people trying to produce viral content to get clicks to direct others to a website full of Google ads," says Giovanni Zagni, director of Facta, a new Italian fact-checking website. Zagni says the site has focused about 90% of its content on COVID-19 since its launch on 2 April.

> **"Sharing your work and expertise, and engaging with the public, is an important part of being a scientist now."**

Many COVID-19 myths seem to be politically motivated, such as the reports that SARS-CoV-2 either escaped from the Wuhan Institute of Virology in China or was a bioweapon created deliberately in the country. A survey of US residents conducted in mid-March found that 6% thought the virus was accidentally created in a laboratory, and 23% that it was developed intentionally (see go.nature.com/2zf4v4d).

Scientists might have more impact when confronting myths that are less political. "If it's something crazy, like the virus is a bioweapon created by Barack Obama, I think scientists are better off leaving that to others and spending their time in the world of science," says West. Scientists can offer their expertise to journalists and fact-checkers who are debunking misinformation.

But should scientists attempt to counter misinformation across fields, or stick to their own? The debate over whether researchers should 'stay in their lane' has, at times, become heated during the COVID-19 pandemic.

In March, the UK-based Science Media Centre, which provides journalists with comments and briefings from scientists, asked its network of experts to stick to their disciplines when responding to media queries about COVID-19. Others, such as West, disagree. "We should encourage, not discourage, scientists to 'step outside their lane', especially during a worldwide crisis," he says. "As long as they are transparent about their expertise, there is much to gain from more scientists thinking about the problem."

Friendly fire

The tone of interventions can determine how they are received. In March, British singer and television personality Kerry Katona shared an Instagram post claiming that children with COVID-19 would be separated from their parents and taken to hospital alone. British doctor and television presenter Ranjit Singh responded: "Not true! Facts are facts! I've seen lots of confusion & misinformation about kids & #coronavirus recently," and posted a summary of the correct information from the UK Royal College of Paediatrics & Child Health. Katona thanked him and said she felt reassured. Zagni says that avoiding appearing confrontational or patronizing is key when seeking to change minds.

Subtle reminders about accuracy that avoid direct confrontation might prove effective. In a study currently awaiting peer review (G. Pennycook *et al.* Preprint at https://psyarxiv.com/uhbk9/; 2020), psychologist Gordon Pennycook, at the University of Regina in Canada, showed two groups of people from the United States a series of news headlines about COVID-19. Half of the headlines were true and half were false; the participants were not told which was which. On average in the first group, 47% of the accurate headlines and 43% of the inaccurate ones were considered worth sharing. The second group was asked to rate the accuracy of a single headline unrelated to COVID-19 before performing the same task. This seemed to make them more discerning, because they went on to say they would consider sharing 50% of the true reports and 40% of the untrue ones.

Many of those who have been inspired to use their training and experience as scientists to protect people from false information about COVID-19 simply want to contribute to reducing the loss of life and health. There could, however, be other benefits to getting involved in the defence of scientific truth. "Sharing your work and expertise, and engaging with the public, is an important part of being a scientist now," says Samantha Vanderslott, a health sociologist at the University of Oxford, UK. "Calling out fake stories can raise your profile."

EIGHT WAYS TO SPOT MISINFORMATION

Health sociologist Samantha Vanderslott at the University of Oxford, UK, studies how ideas, including misinformation, are spread through social media as part of her work on parental attitudes and decisions about vaccination. Here are her top tips on how to boost your immunity to online myths, lies, scams and hoaxes.

Source suspicion. Vague, untraceable sources, such as 'a doctor friend of a friend' or 'scientists say' without further details, should ring alarm bells.

Bad language. Most trustworthy sources are regular communicators, so poor spelling, grammar or punctuation are grounds for suspicion.

Emotional contagion. If something makes you angry or overjoyed, be on your guard. Miscreants know that messages that trigger strong emotions get shared the most.

News gold or fool's gold? Genuine scoops are rare. If information is reported by only one source, beware — especially if it suggests that something is being hidden from you.

False accounting. Use of fake social-media accounts, such as @BBCNewsTonight, is a classic trick. Look out for misleading images and bogus web addresses, too.

Oversharing. If someone urges you to share their sensational news, they might just want a share of the resulting advertising revenue.

Follow the money. Think about who stands to gain from you believing extraordinary claims.

Fact-check check. Go past the headlines and read a story to the end. If it sounds dubious, search fact-checking websites to see whether it has already been debunked.

Overall, West argues that researchers shouldn't allow professional considerations to get in the way when deciding whether to help in the battle against COVID-19 misinformation. "Ultimately, it really shouldn't matter, because lives, and trust in science, are at stake and we need to do something about it."

Nic Fleming is a science writer based in Bristol, UK.

ILLUSTRATION BY THE PROJECT TWINS

WHAT PROJECT-MANAGEMENT SOFTWARE CAN DO FOR SCIENTISTS

Four tools that help researchers to see the big picture, especially when working in collaborative groups. **By Anna Nowogrodzki**

In January 2019, NASA announced that its Transiting Exoplanet Survey Satellite had discovered a planet about three times the diameter of Earth. The planet, orbiting a dwarf star 16 parsecs (53 light years) away, was found using sophisticated equipment including the satellite itself and the Magellan II telescope at Las Campanas Observatory in Chile. But its discovery also relied on a more prosaic tool, says astronomer Johanna Teske: the project-management software Trello.

The five-university consortium that oversees the telescope uses Trello to track and manage the queue of astronomical targets that different teams want to observe, says Teske, who works at Carnegie Observatories in Pasadena, California. "The way that Trello organizes information seemed very much in line with the type of information we wanted to capture," she says, and it's worked well.

Popular project-management tools for research teams include Trello and Jira, both from the company Atlassian in Sydney, Australia, as well as Asana and GitHub project boards, both in San Francisco, California. These tools are more than simple to-do lists. They help teams to see the broad view of a project, allowing users to create and complete tasks, meet deadlines, capture detail-rich notes and provide templates for common protocols. The tagging functions of these tools allow managers to assign tasks to team members. If used well, they can make teams more efficient and minimize frustrations such as forgotten tasks and duplicated work.

In short, project-management tools and the managers who use them "connect the details with the high-level goals", says Tracy Teal. As the executive director of Dryad, a non-profit repository for open data in Durham, North Carolina, she uses several such tools.

Management experience

At the Broad Institute of Harvard and MIT in Cambridge, Massachusetts, computational biologist Beth Cimini manages a small consultancy group within a larger team run by cell biologist Anne Carpenter, which focuses on automated image analysis. Carpenter's group uses project-management tools to handle tasks ranging from keeping track of the team's overall direction to experimental design, Cimini says — the latter thanks to a Trello template that automatically pre-populates notes with standard operating procedures so that laboratory members don't forget key steps. "It's definitely reduced the amount of time we spend reproducing what someone else has already done," Cimini explains. Her own team uses Trello and GitHub project boards to juggle their clients' needs. "It would be hard for each person in our group to have ten different projects a year" without them, she says.

Project-management tools tend to have a common visual style, called a Kanban board. This is divided into columns, called lists, with multiple cards pinned to them to represent different projects, protocols or topics. Users can make multiple lists (for example, 'To do', 'In progress' and 'Done'), create individual to-do items (either in the app or by sending an e-mail to a dedicated account), tag teammates to assign tasks, and drag the cards from board to board as their status changes. Many tools can also display a timeline or calendar view, and provide apps for use on mobile devices.

But there are differences, and most tools offer both free and paid tiers, with different incentives for paid accounts. "It's worth exploring the different tools a bit and finding the right ones for you," recommends José Sánchez-Gallego, an astronomer at the University of Washington in Seattle. "Personally, I prefer tools that do one thing but do it well, rather than tools that allow you to do many things but become more cumbersome."

Sánchez-Gallego actually uses multiple tools for project management in his day-to-day work. These include ZenHub for managing GitHub issues for the Sloan Digital Sky Survey telescope in New Mexico; Jira for overarching project management, hardware issues and input from users; and OmniPlan for creating timelines and tracking time. "I like to look for simplicity and good overall design," Sánchez-Gallego says. "I also prefer apps that can work offline over web apps that only work when connected to the Internet. And I prefer tools that don't require me to share too much personal information."

With any project-management tool, the most difficult part is establishing a routine for using it, says Cimini. "It's easier to enforce doing that when it's collaborative," she says. "My collaborative Trello boards stay more up to date sometimes than my personal one."

Appoint a manager to run the tool at a team level, Teal advises. The Data Curation Network, of which Dryad is a partner, has a project manager who goes through the team's Jira to-do items and pings a message to people if tasks aren't done, Teal notes. "The social connection between the tool and the team is a person who consistently makes that connection," she says.

> **"If used well, these tools can make teams more efficient and minimize frustrations."**

Whichever project-management tool you use, ease your team into it to avoid overwhelming them, says Rafael Carazo Salas, who began using Trello after coronavirus-related shutdowns to aid communication and assign tasks in his stem-cell differentiation lab at the University of Bristol, UK. And don't feel you must restrict yourself to tasks, he adds. Salas has started using Trello to share literature with his team, tagging members on articles that are especially relevant to them. The literature keeps Trello interesting, and the tags alert people until checking the tool becomes a habit, he says. "Make it reach out to them," says Salas, "instead of making it a static board that they need to actively go and check."

Project toolbox

For the Magellan II telescope collaboration, Teske says, it is Trello's nested structure that allows the team to manage its users' needs. There's a board for each of the five institutions, which is visible only to that institution, and a separate board for the administrative team that is filled with astronomical targets for each slot of observing time. When a scientist wants telescope time, they create a card on their institution's board, which the administration team then moves to the observation board. Cards can include notes, PDFs and data files, and any other useful information. An archive board serves as a record of everything that has been done. "I think people find it intuitive," Teske says.

But small teams can also benefit from such tools. For Cimini's five-member team, Trello's integration with Tick, its time-tracking and billing software, has proved particularly useful to automatically track the amount of time they've spent working on projects that are billed separately, or to allocate how much time to spend on specific tasks. (Asana also integrates with Tick.)

Pre-formatted boards called templates are also useful, because they provide a starting place for common tasks. In Cimini's group, every time a team member kicks off one of their standard experiments, they use a template so they can be sure of completing every step in the protocol, she says. Cimini has also created a template for travel, which includes standard tasks such as booking flights and hotels, and preparing presentations. This feature is particularly useful, she says, because the travel card stays on her Trello board until she remembers to file for reimbursements.

In her previous position at The Carpentries in Oakland, California, an organization that teaches coding and data workshops, Teal and her co-workers used Asana templates to ensure that they would remember to add essential components such as context, recurring tasks and milestones to every project. And they had a standard template to ensure that they completed all the tasks in the right order to be able to launch their workshops.

Project-management tools typically support plug-ins to enhance functionality. Trello, Jira and Asana can all integrate with the code-sharing site GitHub, for instance. But for developers and scientists who already spend a lot of time on the site, GitHub project boards are particularly appealing, say Teal and Cimini, whose teams both use this tool.

GitHub is a collaborative platform for people who develop software. Project boards organize GitHub's social elements — issue trackers, comments and code updates called pull requests — into a Kanban-like board. "It's this quick graphical way to understand how behind I am," says Clair Sullivan, a machine-learning engineer at GitHub, who is based in Breckenridge, Colorado. Whenever a programmer flags an issue (such as a bug report or a request for a new feature in the software), the software automatically slots it into the board's to-do column. As the team addresses these requests with finalized code fixes, GitHub's built-in Actions tool automatically marks the issues as done.

Sánchez-Gallego spends a lot of time using GitHub when he works with the team that maintains Marvin, an open-source data-visualization tool. But for his work managing the Sloan Digital Sky Survey helpdesk, he favours Jira, which his team has found to be more accessible for people who do not have experience in developing software. Observers and technicians at the two observatories his team supports use Jira to log tickets when something goes wrong. "What I find most useful is the ability to create personal filters," he says. This lets him see only the tickets that are most relevant to him.

No matter which management tool you choose, engage your team early in the decision-making process, Teal advises. Think about their needs and how they spend their time — for example, on GitHub or in their e-mail inboxes. Otherwise, your project-management tool risks becoming "sort of like another inbox", she says — just another thing that's hard to remember to check.

Anna Nowogrodzki is a journalist in Boston, Massachusetts.

Where I work

Claudine Loisel

Photographed by Francois Mori/AP/Shutterstock.

Stained glass is a magical material, whether in a church or a laboratory. Here, I'm examining a panel taken from a French national treasure – Notre-Dame cathedral in Paris – after it was nearly destroyed by fire on 15 April 2019. I'm leaning over a light table at the Historical Monuments Research Laboratory, my workspace west of the city. The lab closed because of the COVID-19 outbreak, but reopened on 3 June.

I'm peering intently at this detail showing the robe of King David, from a nineteenth-century painting by Charles-Laurent Maréchal, whose stained-glass work appears in cathedrals across France. This panel was especially close to the fire, so we wanted to check it for damage.

I'm wearing protective gear to shield myself from possible exposure to lead. The framework holding the glass in place is loaded with this metal, but the risk turned out to be negligible: lead melts at around 328°C, but the windows never got that hot.

I can see some small pathologies in this glass – including a few smooth, rounded cracks that signify thermal shock – but overall we were very lucky. The firefighters did an amazing job. They knew that the windows could explode if they got wet, and they managed to control the blaze without spraying the windows.

As a glass specialist, I study the chemistry of stained glass at a microscopic and nanoscopic scale. I'm fascinated by the materials used and the evolution of techniques. You have to respect the artists. When you see a piece of glass that has barely degraded over hundreds of years, it's almost unbelievable. Glass holds many secrets.

Science aside, the first thing you notice about stained glass is its beauty. I'm very lucky to be in this field. And I'm part of an amazing team of historians, conservators and materials specialists working to restore, protect and eventually reopen Notre-Dame. After that, we'll have a glass of champagne.

Claudine Loisel is a glass specialist at the Historical Monuments Research Laboratory in Champs-Sur-Marne, France. **Interview by Chris Woolston**.